

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2002, 06:11:33 ; Search time 51 Seconds

(without alignments)
78.173 Million cell updates/sec

Title: US-09-548-449-8
Perfect score: 13
Sequence: 1 aatcaannanta 13

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 687286

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCMTUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10	76.9	29	4	US-09-593-012-112
2	10	76.9	36	2	US-08-669-721-1
3	10	76.9	36	2	US-08-669-721-1
4	10	76.9	36	4	US-09-189-344-1
5	10	76.9	36	4	US-09-189-344-1
6	10	76.9	96	6	5190931-4
7	10	76.9	100	1	US-08-145-705A-17
8	9.2	70.8	72	1	US-07-737-071A-3
9	9.2	70.8	72	1	US-08-022-096-3
10	9	69.2	15	1	US-08-334-847-467
11	9	69.2	17	3	US-08-985-162-812
12	9	69.2	17	3	US-08-985-162-812
13	9	69.2	17	3	US-08-985-162-812
14	9	69.2	18	4	US-09-241-581B-2
15	9	69.2	18	4	US-08-265-428-2
16	9	69.2	18	5	PCT-US95-07721-2
17	9	69.2	20	2	US-08-743-637B-39
18	9	69.2	20	3	US-08-526-840B-39
19	9	69.2	20	3	US-08-680-506-18
20	9	69.2	20	3	US-09-357-073-43
21	9	69.2	20	3	US-09-288-461-20
22	9	69.2	20	4	US-09-889-595-17
23	9	69.2	21	1	US-08-328-961-12
24	9	69.2	21	1	US-08-462-397-12
25	9	69.2	21	2	US-08-468-819-30
26	9	69.2	21	2	US-08-468-819-30
27	9	69.2	22	2	US-08-613-417A-25

28	9	69.2	22	3	US-08-594-452-25	Sequence 25, Appl
29	9	69.2	22	3	US-08-578-686C-24	Sequence 24, Appl
30	9	69.2	22	3	US-08-281-203-18	Sequence 18, Appl
31	9	69.2	22	3	US-09-094-405-29	Sequence 29, Appl
32	9	69.2	22	3	US-09-258-408-25	Sequence 25, Appl
33	9	69.2	22	3	US-09-196-132-25	Sequence 25, Appl
34	9	69.2	22	3	US-09-144-112-24	Sequence 24, Appl
35	9	69.2	22	4	US-08-337-120A-33	Sequence 33, Appl
36	9	69.2	24	1	US-07-722-798A-74	Sequence 74, Appl
37	9	69.2	24	4	US-09-561-825-12	Sequence 12, Appl
38	9	69.2	24	4	US-09-514-302-10	Sequence 10, Appl
39	9	69.2	26	1	US-07-722-798A-73	Sequence 73, Appl
40	9	69.2	26	3	US-08-594-452-97	Sequence 97, Appl
41	9	69.2	26	3	US-08-594-452-98	Sequence 98, Appl
42	9	69.2	26	3	US-09-258-408-97	Sequence 97, Appl
43	9	69.2	26	4	US-09-258-408-98	Sequence 98, Appl
44	9	69.2	29	4	US-09-302-620B-8	Sequence 8, Appl
45	9	69.2	30	1	US-08-460-784-72	Sequence 72, Appl

ALIGNMENTS

RESULT 1
US-09-593-012-112/c
Sequence 112, Application US/09593012
Patent No. 6387652
GENERAL INFORMATION:
APPLICANT: HAUGLAND, Richard
APPLICANT: VESPER, Stephen
TITLE OF INVENTION: METHOD OF IDENTIFYING AND QUANTIFYING SPECIFIC FUNGI AND BACTE
FILE REFERENCE: HAUGLAND-1A
CURRENT APPLICATION NUMBER: US/09/593,012
CURRENT FILING DATE: 2000-06-13
PRIOR APPLICATION NUMBER: US 09/290,990
PRIOR FILING DATE: 1999-04-14
PRIOR APPLICATION NUMBER: US 60/081,773
PRIOR FILING DATE: 1998-04-15
NUMBER OF SEQ ID NOS: 225
SOFTWARE: PatentIn version 3.0
SEQ ID NO 112
LENGTH: 29
TYPE: DNA
ORGANISM: Paecilomyces variotii
US-09-593-012-112

Query Match 76.9%; Score 10; DB 4; Length 29;
Best Local Similarity 76.9%; Pred. No. 1.6e+03;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 AATCAANNANTTA 13
Db 27 AATCAATCAATTA 15

RESULT 2
US-08-669-721-1
Sequence 1, Application US/08669721
Patent No. 5834236
GENERAL INFORMATION:
APPLICANT: Lamb et al., Christopher J.
TITLE OF INVENTION: NOVEL TRANSCRIPTION ENHANCER ELEMENT AND
TITLE OF INVENTION: TRANSCRIPTION FACTOR AND METHODS OF USE THEREFOR
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/669,721
FILING DATE: 27-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ellison, Eldora L.
REGISTRATION NUMBER: 39,967
REFERENCE/DOCKET NUMBER: 07251/014001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-669-721-1

Query Match 76.9%; Score 10; DB 2; Length 36;
Best Local Similarity 76.9%; Pred. No. 1.6e+03;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 AATCAANNNTTA 13
|||||
DB 9 AATCAATTAATTA 21

RESULT 3
US-08-669-721-1/c
Sequence 1, Application US/08669721
Patent No. 5834236
GENERAL INFORMATION:
APPLICANT: Lamb et al., Christopher J.
TITLE OF INVENTION: NOVEL TRANSCRIPTION ENHANCER ELEMENT AND
TITLE OF INVENTION: TRANSCRIPTION FACTOR AND METHODS OF USE THEREFOR
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/669,721
FILING DATE: 27-JUN-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ellison, Eldora L.
REGISTRATION NUMBER: 39,967
REFERENCE/DOCKET NUMBER: 07251/014001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-669-721-1

Query Match 76.9%; Score 10; DB 2; Length 36;
Best Local Similarity 76.9%; Pred. No. 1.6e+03;

Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 AATCAANNNTTA 13
|||||
DB 32 AATCAATTAATTA 20

RESULT 4
US-09-189-344-1
Sequence 1, Application US/09189344
Patent No. 6191258
GENERAL INFORMATION:
APPLICANT: Lamb et al., Christopher J.
TITLE OF INVENTION: NOVEL TRANSCRIPTION ENHANCER ELEMENT AND
TITLE OF INVENTION: TRANSCRIPTION FACTOR AND METHODS OF USE THEREFOR
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/189,344
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/669,721
FILING DATE: 27-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ellison, Eldora L.
REGISTRATION NUMBER: 39,967
REFERENCE/DOCKET NUMBER: 07251/014001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-189-344-1

Query Match 76.9%; Score 10; DB 4; Length 36;
Best Local Similarity 76.9%; Pred. No. 1.6e+03;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 AATCAANNNTTA 13
|||||
DB 9 AATCAATTAATTA 21

RESULT 5
US-09-189-344-1/c
Sequence 1, Application US/09189344
Patent No. 6191258
GENERAL INFORMATION:
APPLICANT: Lamb et al., Christopher J.
TITLE OF INVENTION: NOVEL TRANSCRIPTION ENHANCER ELEMENT AND
TITLE OF INVENTION: TRANSCRIPTION FACTOR AND METHODS OF USE THEREFOR
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA

COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/189,344
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/669,721
FILING DATE: 27-JUN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ellison, Eldora L.
REGISTRATION NUMBER: 39,967
REFERENCE/DOCKET NUMBER: 07251/014001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-189-344-1

Query Match 76.9%; Score 10; DB 4; Length 36;
Best Local Similarity 76.9%; Pred. No. 1.5e+03;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AATCAANNANTTA 13
||||| | |||
Db 32 AATCAATTAATTA 20

RESULT 6
5190931-4
PATENT NO. 5190931
APPLICANT: INOUE, MASAYORI
TITLE OF INVENTION: REGULATION OF GENE EXPRESSION BY
EMPLOYING TRANSLATIONAL INHIBITION OF mRNA UTILIZING
INTERFERING COMPLEMENTARY RNA
NUMBER OF SEQUENCES: 7
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/436,598
FILING DATE: 15-NOV-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 300,741
FILING DATE: 23-JAN-1989
APPLICATION NUMBER: 228,852
FILING DATE: 03-AUG-1988
APPLICATION NUMBER: 543,528
FILING DATE: 20-OCT-1983
SEQ ID NO: 4:
LENGTH: 96
5190931-4

Query Match 76.9%; Score 10; DB 6; Length 96;
Best Local Similarity 76.9%; Pred. No. 1.5e+03;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AATCAANNANTTA 13
||||| | |||
Db 42 AATCAATTAATTA 54

RESULT 7
US-08-145-705A-17/C
Sequence 17, Application US/08145705A
Patent No. 5489513

GENERAL INFORMATION:
APPLICANT: Springer, Wolfgang, Plempel, Manfred;
APPLICANT: L. Bieding, Antonius
TITLE OF INVENTION: SPECIFIC GENE PROBES AND
TITLE OF INVENTION: PROCESSES FOR THE DIAGNOSTIC
TITLE OF INVENTION: INVESTIGATION OF CANDIDA
TITLE OF INVENTION: ALBICANS
NUMBER OF SEQUENCES: 44
CORRESPONDENCE ADDRESS:
ADDRESSEE: SPRUNG HORN KRAMER & WOODS
STREET: 660 White Plains Road
CITY: Tarrytown
STATE: New York
COUNTRY: U.S.A.
ZIP: 10591-5144
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.4 MB storage
COMPUTER: NEC Powermate 1 Plus
OPERATING SYSTEM: DOS
SOFTWARE: WordPerfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/145,705A
FILING DATE: October 28, 1993
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: German P 42 36 708.5
FILING DATE: October 30, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Kurt G. Briscoe
REGISTRATION NUMBER: 33,141
REFERENCE/DOCKET NUMBER: Bayer 8885-KGB
TELECOMMUNICATION INFORMATION:
TELEPHONE: (914) 332-1700
TELEFAX: (914) 332-1844
TELEX:
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 100 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Candida albicans
US-08-145-705A-17

Query Match 76.9%; Score 10; DB 1; Length 100;
Best Local Similarity 76.9%; Pred. No. 1.5e+03;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AATCAANNANTTA 13
||||| | |||
Db 56 AATCAATTAATTA 44

RESULT 8
US-07-737-071A-3/C
Sequence 3, Application US/07737071A
Patent No. 5229286
GENERAL INFORMATION:
APPLICANT: JARSCHE, Michael
APPLICANT: LANG, Gunter
TITLE OF INVENTION: CLONING AND OVEREXPRESSION OF
TITLE OF INVENTION: GLUCOSE-6-PHOSPHATE DEHYDROGENASE FROM LEUCONOSTOC
TITLE OF INVENTION: DEXTRANSUC
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Armstrong, Nikaido, Marmelstein Kubovick &
ADDRESSEE: Murray
STREET: 1725 K Street, N.W., Suite 1000
CITY: Washington D.C.

```

1 COUNTRY: United States of America
2 ZIP: 20006
3
4 COMPUTER READABLE FORM:
5 MEDIUM TYPE: floppy disk
6 COMPUTER: IBM PC compatible
7 OPERATING SYSTEM: PC-DOS/MS-DOS
8 SOFTWARE: Patent Release #1.0, Version #1.25
9 CURRENT APPLICATION DATA:
10 APPLICATION NUMBER: US/07/737,071A
11 FILING DATE: 19910730
12 CLASSIFICATION: 435
13 PRIOR APPLICATION DATA:
14 APPLICATION NUMBER: DE P 4024158.9
15 FILING DATE: 30-JUL-1990
16 ATTORNEY/AGENT INFORMATION:
17 NAME: Murray, Robert B.
18 REGISTRATION NUMBER: 22,890
19 REFERENCE/DOCKET NUMBER: 911028
20 TELECOMMUNICATION INFORMATION:
21 TELEPHONE: (202) 659-2930
22 TELEFAX: (202) 887-0357
23 TELEX: 440142
24 INFORMATION FOR SEQ ID NO: 3:
25 SEQUENCE CHARACTERISTICS:
26 LENGTH: 72 base pairs
27 TYPE: NUCLEIC ACID
28 STRANDEDNESS: single
29 TOPOLOGY: linear
30
31 US-07-737-071A-3
32
33 Query Match 70.8%; Score 9.2; DB 1; Length 72;
34 Best Local Similarity 61.5%; Pred. No. 4.4e+03;
35 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0.
36
37 QY 1 AATCAANNATTA 13
38 ||||| |||
39 db 51 AATCAATTATTM 39
40
41 RESULT 9
42 US-08-022-096-3/c
43 Sequence 3, Application US/08022096
44 Patent No. 5308770
45
46 GENERAL INFORMATION:
47 APPLICANT: JARSCH, Michael
48 APPLICANT: LANG, Gunter
49 TITLE OF INVENTION: CLONING AND OVEREXPRESSION OF
50 TITLE OF INVENTION: GLUCOSE-6-PHOSPHATE DEHYDROGENASE FROM LEUCONOSTOC
51 TITLE OF INVENTION: DETRANICUS
52 NUMBER OF SEQUENCES: 5
53 CORRESPONDENCE ADDRESS:
54 ADDRESSEE: Nikaido, Marmelstein, Murray & Oram
55 STREET: 655 Fifteenth Street N.W. Suite 330
56 CITY: Washington
57 STATE: D.C.
58 COUNTRY: U.S.A.
59 ZIP: 20005-5701
60
61 COMPUTER READABLE FORM:
62 MEDIUM TYPE: floppy disk
63 COMPUTER: IBM PC compatible
64 OPERATING SYSTEM: PC-DOS/MS-DOS
65 SOFTWARE: Patent Release #1.0, Version #1.25
66 CURRENT APPLICATION DATA:
67 APPLICATION NUMBER: US/08/022,096
68 FILING DATE: 19930225
69 CLASSIFICATION: 435
70 PRIOR APPLICATION DATA:
71 APPLICATION NUMBER: DE P 40 24 158.0
72 FILING DATE: 30-JUL-1990
73 PRIOR APPLICATION DATA:
74 APPLICATION NUMBER: US 07/737,071
75 FILING DATE: 30-JUL-1991
76 ATTORNEY/AGENT INFORMATION:

```

NAME: Chin, Monica F.
REGISTRATION NUMBER: 36, 105
REFERENCE/DOCKET NUMBER: P564-3005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)638-5000
TELEFAX: (202)638-4810
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 72 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
US-08-022-096-3

Query Match
Best Local Similarity 70.8%; Score 9.2; DB 1; Length 72;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 51 AATCAATTTT 39

RESULT 40
US-08-334-847-467
Sequence 467, Application US/08334847
Patent No. 5693532
GENERAL INFORMATION:
APPLICANT: MCSwigen, James
APPLICANT: Draper, Kenneth
APPLICANT: Payco, Paul
APPLICANT: Woolf, Tod
TITLE OF INVENTION: METHOD AND REAGENT FOR
TITLE OF INVENTION: INHIBITING RESPIRATORY
TITLE OF INVENTION: SYNCYTIAL VIRUS
NUMBER OF SEQUENCES: 909
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1.
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/334, 847
FILING DATE: No. 5693532member 4, 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Waiburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 209/032
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 467:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-334-847-467

Query Match 69.2%; Score 9; DB 1; Length 15;

Best Local Similarity 50.0%; Pred. No. 5.9e+03;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
OY 2 ATCAANNANTT 13
|:|:| | : : |
Db 2 AUCAAUAGUUA 13

RESULT 11
US-08-985-162-812/C

; Sequence 812, Application US/08985162
; Patent No. 6057156
; GENERAL INFORMATION:
; APPLICANT: Akhtar, Saghir
; APPLICANT: McSwiggen, James
; TITLE OF INVENTION: ENZYMATIC NUCLEIC ACID TREATMENT
; TITLE OF INVENTION: OF DISEASES OR CONDITIONS RELATED
; TITLE OF INVENTION: TO LEVELS OF EPIDERMAL GROWTH
; NUMBER OF SEQUENCES: 1877
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,162
; FILING DATE: 04 December 1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/036,476
; FILING DATE: 31 January 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 230/107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 812:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-985-162-812

Query Match 69.2%; Score 9; DB 3; Length 17;
Best Local Similarity 75.0%; Pred. No. 5.9e+03;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 AATCAANNANTT 12
|:|:| | | : |
Db 16 AATCAAAATCTT 5

RESULT 12
US-08-985-162-813/C
; Sequence 813, Application US/08985162
; Patent No. 6057156
; GENERAL INFORMATION:
; APPLICANT: Akhtar, Saghir

; APPLICANT: Fell, Patricia
; APPLICANT: McSwiggen, James
; TITLE OF INVENTION: ENZYMATIC NUCLEIC ACID TREATMENT
; TITLE OF INVENTION: OF DISEASES OR CONDITIONS RELATED
; TITLE OF INVENTION: TO LEVELS OF EPIDERMAL GROWTH
; NUMBER OF SEQUENCES: 1877
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq for Windows 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/985,162
; FILING DATE: 04 December 1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/036,476
; FILING DATE: 31 January 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 230/107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 813:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-985-162-813

Query Match 69.2%; Score 9; DB 3; Length 17;
Best Local Similarity 75.0%; Pred. No. 5.9e+03;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 1 AATCAANNANTT 12
|:|:| | | : |
Db 14 AATCAAAATCTT 3

RESULT 13
US-08-985-162-814/C
; Sequence 814, Application US/08985162
; Patent No. 6057156
; GENERAL INFORMATION:
; APPLICANT: Akhtar, Saghir
; APPLICANT: McSwiggen, James
; TITLE OF INVENTION: ENZYMATIC NUCLEIC ACID TREATMENT
; TITLE OF INVENTION: OF DISEASES OR CONDITIONS RELATED
; TITLE OF INVENTION: TO LEVELS OF EPIDERMAL GROWTH
; NUMBER OF SEQUENCES: 1877
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.

ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/985,162
FILING DATE: 04 December 1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/036,476
FILING DATE: 31 January 1997
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 230/107
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 814:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-985-162-814

Query Match 69.2%; Score 9; DB 3; Length 17;
Best Local Similarity 75.0%; Pred. No. 5.9e+03;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AATCAANNNTT 12
||||| ||
DB 13 AATCAATACT 2

RESULT 14
US-09-241-581B-2
Sequence 2, Application US/09241581B
Patent No. 6350859
GENERAL INFORMATION:
APPLICANT: Massachusetts Institute of Technology
TITLE OF INVENTION: Class BI and CI Scavenger Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 2800 One Atlantic Center
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/241,581B
FILING DATE: 02-Feb-1999
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: MIT6620
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

LENGTH: 18 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
PUBLICATION INFORMATION:
AUTHORS: Ashkenas, et al.
JOURNAL: J. Lipid Res.
VOLUME: 34
PAGES: 983-1000
DATE: 1993
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 18
US-09-241-581B-2

Query Match 69.2%; Score 9; DB 4; Length 18;
Best Local Similarity 75.0%; Pred. No. 5.9e+03;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AATCAANNNTT 12
||||| ||
DB 1 AATCAAGAAAT 12

RESULT 15
US-08-265-428-2
Sequence 2, Application US/08265428
Patent No. 6429289
GENERAL INFORMATION:
APPLICANT: Krieger, Monty
TITLE OF INVENTION: Class BI Scavenger Receptors
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Patrea L. Pabst
STREET: 1100 Peachtree Street, Suite 2800
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30309-4530
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/265,428
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: MIT6620
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 815-6558
TELEFAX: (404) 815-6555
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
PUBLICATION INFORMATION:
AUTHORS: Ashkenas, et al.
JOURNAL: J. Lipid Res.
VOLUME: 34
PAGES: 983-1000
DATE: 1993
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 18

US-08-265-428-2

Query Match 69.28; Score 9; DB 4; Length 18;
 Best Local Similarity 75.08; Pred. No. 5.9e+03;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AATCANNANTT 12
 ||||| | ||
 Db 1 AATCAGGAATT 12

Search completed: December 26, 2002, 11:37:15
 Job time : 57 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 25, 2002, 20:55:28 ; Search time 2628 Seconds

(without alignments)
5448.470 Million cell updates/sec

Title: US-09-548-449-3

Perfect score: 492

Sequence: 1 actcgcgacatctacc.....gccttaacctagcaaggac 492

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 100%

Listing first 45 summaries

Database : GenBml:*

1: gb_ba:*

2: gb_hlg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vl:*

30: em_hlg_hum:*

31: em_hlg_inv:*

32: em_hlg_other:*

33: em_hlg_mus:*

34: em_hlg_pln:*

35: em_hlg_rod:*

36: em_hlg_mam:*

37: em_hlg_vrt:*

38: em_sy:*

39: em_hlgo_hum:*

40: em_hlgo_mus:*

41: em_hlgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	492	100.0	3761	1 PANR	X57736 P.aeruginos
2	492	100.0	3761	1 PSEANRPRO	M98276 Pseudomonas
3	490.4	99.7	11627	1 AE004582	AE004582 Pseudomon
4	281	57.1	2635	1 PAHEMN	X97961 P.aeruginos
5	90	18.3	900	1 PANR	X58405 P.aeruginos
6	51.8	10.5	129420	8 AC087547	AC087547 Oryza sat
7	51.8	10.5	141041	8 AC087550	AC087550 Oryza sat
8	49.8	10.1	175644	2 AP003771	AP003771 Oryza sat
9	49.4	10.0	68736	2 AC090435	AC090435 Chlamydom
10	49.4	10.0	102657	2 AC090436	AC090436 Chlamydom
11	49.2	10.0	3312	1 AF061246	AF061246 Ralstonia
12	48.2	9.8	2450	1 AF033262	AF033262 Pseudomon
13	48	9.8	2668	9 AF355797	AF355797 Homo sapi
14	48	9.8	38939	9 AC004678	AC004678 Homo sapi
15	47.8	9.7	13076	9 HSBRN3A2	U10063 Human POU d
16	47.6	9.7	13076	1 AE012233	AE012233 Xanthomon
17	47.6	9.7	40352	1 SC661A	AL356595 Streptomy
18	47.2	9.6	57078	2 AC125442	AC125442 Mus muscu
19	47.2	9.6	165497	8 AP003232	AP003232 Oryza sat
20	47	9.6	138904	2 AP003928	AP003928 Oryza sat
21	47	9.6	153180	2 AP004683	AP004683 Oryza sat
22	46.8	9.5	123580	1 AF263912	AF263912 Streptomy
23	46.8	9.5	125401	6 AX211739	AX211739 Sequence
24	46.8	9.5	134367	8 AP003842	AP003842 Oryza sat
25	46.8	9.5	169162	8 AP004267	AP004267 Oryza sat
26	46.4	9.4	42348	1 SC27611	AL391763 Streptomy
27	46.2	9.4	123071	2 AC121481	AC121481 Rattus no
28	46	9.3	2765	6 BD012188	BD012188 Rattus no
29	46	9.3	2765	23 BD004868	BD004868 Novel can
30	46	9.3	3131	6 BD012187	BD012187 Novel can
31	46	9.3	3131	23 BD004867	BD004867 Novel can
32	46	9.3	3160	9 AB043584	AB043584 Homo sapi
33	46	9.3	3332	6 BD012186	BD012186 Novel can
34	46	9.3	3332	23 BD004866	BD004866 Novel can
35	46	9.3	3344	6 BD012182	BD012182 Novel can
36	46	9.3	3344	23 BD004862	BD004862 Novel can
37	46	9.3	3842	1 AF106004	AF106004 Streptomy
38	46	9.3	7603	9 HSA404614	AJ404614 Homo sapi
39	46	9.3	10206	1 AE012548	AE012548 Xanthomon
40	46	9.3	14478	1 SC2D4	AL392175 Streptomy
41	46	9.3	31624	1 SCDB3	AL161755 Streptomy
42	46	9.3	42526	1 SC4AN7	AL133423 Streptomy
43	46	9.3	58930	2 AC098321	AC098321 Rattus no
44	46	9.3	104782	9 CNS05TEZ	AL359681 Human chr
45	46	9.3	124539	2 AC044794	AC044794 Homo sapi

ALIGNMENTS

RESULT 1

PANR 3761 bp DNA linear BCT 18-NOV-1997

LOCUS X57736

DEFINITION P.aeruginosa anr gene for positive control element ANR.

ACCESSION X57736

VERSION X57736.1 GI:45359

KEYWORDS anr gene; ANR protein; positive control element.

SOURCE Pseudomonas aeruginosa.

ORGANISM Pseudomonas aeruginosa.

REFERENCE 1 (bases 1 to 3761)

AUTHORS Zimmermann, A.

TITLE Direct Submission

JOURNAL Submitted (12-FEB-1991) A. Zimmermann, Institut f Mikrobiologie ETH

REMARK
Zuerich, Schmelzbergstr 7, CH-8092 Zuerich, Switzerland
revised by [3]
2 (bases 1931 to 2958)
REFERENCE
AUTHORS
TITLE
Anaerobic growth and cyanide synthesis of *Pseudomonas aeruginosa*
depend on *aur*, a regulatory gene homologous with *fur* of *Escherichia coli*
JOURNAL
MOL. MICROBIOL. 5 (6), 1483-1490 (1991)
MEDLINE
92157874
PUBMED
1787798
REFERENCE
3 (bases 1 to 3761)
AUTHORS
Zimmermann, A.
TITLE
Direct Submission
JOURNAL
Submitted (25-FEB-1992) A. Zimmermann, same address
FEATURES
Location/Qualifiers
1. 3761
/organism="Pseudomonas aeruginosa"
/strain="PAO1"
/db_xref="taxon:287"
/map="60 min."
/clone_1b="PLAFR3"
22. 504
/note="ORF A"
/codon_start=1
/transl_table=1
/protein_id="CAA40903.1"
/db_xref="GI:45360"
/db_xref="SWISS-PROT:Q04626"
/translation="MLDTRMDADLRIRYDLSGPRYSYPTAVOFHEGIGPEIDHALL
RDSRAGHPLSLVYHIPCANICVYCNKVTITDRGSAPYLARLVEIEIVSHLS
RAOVYBOLHREGGTPFLSPQRLKRLMSQLKTHNLDDSDGDIETIDPEADWSTK
564. 1061
/note="ORF B"
/codon_start=1
/transl_table=1
/protein_id="CAA40904.1"
/db_xref="GI:45361"
/db_xref="SWISS-PROT:Q04627"
/translation="MPQRINADLPSPGQKLEMLQRTTEQLAAAGYRIYIGMDHFLP
DDELASOEDGTLOIRFGYTGHCIDGVAISQIGLYSONSSDIIDYQSLD
NGOLAIRGLHCNSDNRRAVIOQLICHPFLAIEDIETERGIDPRSFALMDLPE
SPPTA"
complement(1235..1702)
/note="ORF X"
/codon_start=1
/transl_table=1
/protein_id="CAA40905.1"
/db_xref="GI:45362"
/db_xref="SWISS-PROT:Q04628"
/translation="MLTSGWITTVQVPPSSRAADAKAARTGTAEQAEPPAGTDAAQ
QHVLDSAGRELAITMGVPEDEDDIPQELRPVVKMIRELKKIEELRELOEMRS
SDPGAKARVPELOKELQOLNSALQATATAAASAIKEMGISDPALIMKMGSR"
1948. 2682
/gene="aur"
1948. 2682
/gene="aur"
/codon_start=1
/transl_table=1
/product="Aur"
/protein_id="CAA40906.1"
/db_xref="GI:45363"
/db_xref="SWISS-PROT:P23926"
/translation="MAETIKYRALPOAHCKDCSLAPLCLPLSTVEDMDSIDEYKRG
RLPKGEFLFGDPGPGSVFAVRSGALKTFSITDAGEBOITGFHLPSLVLSGMDTE
TYPVSAOLETTSVCEIIPERLDSLOLRLROILAMISRETRIDDOOMLLSKTK
ADRIATFLVLSNRFRARFGSAOQFRLAMSRLNIGTLGLAVETYSVFTTRPOOGL
ISAEGKEVHLITELCALAGGLEG"
2768. 3316
/note="ORF C"
/codon_start=1
/transl_table=1
/protein_id="CAA40907.1"

CDS
/db_xref="GI:45364"
/db_xref="SWISS-PROT:Q04633"
/translation="MIFDELTKSQIRAVPDPKPGVYFRDITPLFOSPRLRMTVDS
FVORIEADFSHIGMDARGLIGSAVAVYALNPKPLVERKGGKIPADVLAGYOTEYG
EAFLEVHDSICEGDSVILPDDLATGTGTLAASLVRLGARFEMAATIDPELGG
STRLODAGISTFSLTAPALDER"
3556. 3761
/note="ORF D"
/codon_start=1
/transl_table=1
/protein_id="CAA40908.1"
/db_xref="GI:45365"
/db_xref="SPTRMBL:Q04375"
/translation="MTATGASPVANPVRPMDFSESETPKRYMDQPFMTQPMNLL
SLFPYGEKFEVDSYRAVDRLEDP"
BASE COUNT 669 a 1337 c 1060 g 695 t
ORIGIN
Query Match 100.0%; Score 492; DB 1; Length 3761;
Best Local Similarity 100.0%; Pred. No. 1.8e-73;
Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACTCGGGATCATCTTCACATCGCGCCGCACTCTCGGGATATCTCTCTCTCT 60
DB 1450 ACTCGGGATCATCTTCACATCGCGCCGCACTCTCGGGATATCTCTCTCTCT 1509
QY 61 CCACGGGACACCCCATAGTAGGCGGAGCTCGGCTCTCGGGAACCTGTACATGCT 120
DB 1510 CCACGGGACACCCCATAGTAGGCGGAGCTCGGCTCTCGGGAACCTGTACATGCT 1569
QY 121 GATCGGGGCGGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
DB 1570 GATCGGGGCGGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1629
QY 181 GGGCTTGGCGGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
DB 1630 GGGCTTGGCGGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1689
QY 241 TGAGGGTCAACATTCACATCTCGGGAACATGGAATTCATTCATTCATTCATTC 300
DB 1690 TGAGGGTCAACATTCACATCTCGGGAACATGGAATTCATTCATTCATTCATTC 1749
QY 301 GCGTGGCGCACTTGACCCCTTTTCGTCGCGGCTTGACAGGCTGCGAGCTAGTCGA 360
DB 1750 GCGTGGCGCACTTGACCCCTTTTCGTCGCGGCTTGACAGGCTGCGAGCTAGTCGA 1809
QY 361 GTTGTTCGCGCACTTGATGGAACGCACTGCGGCTGCAAAATGCTGCTTGC 420
DB 1810 GTTGTTCGCGCACTTGATGGAACGCACTGCGGCTGCAAAATGCTGCTTGC 1869
QY 421 AGACCTATGCGTGGACCGGCACTGCGGCTGCTTACCTTACTCTCTTGTGCTTTAA 480
DB 1870 AGACCTATGCGTGGACCGGCACTGCGGCTGCTTACCTTACTCTCTTGTGCTTTAA 1929
QY 481 CCTAGCAAGGAC 492
DB 1930 CCTAGCAAGGAC 1941
RESULT 2
PSEANRPRO 3761 bp DNA linear BCT 26-APR-1993
LOCUS Pseudomonas aeruginosa anaerobic regulatory protein (aur) gene,
DEFINITION complete cds.
ACCESSION M98276
VERSION M98276.1 GI:151019
KEYWORDS aur gene; positive control element; regulatory protein.
SOURCE Pseudomonas aeruginosa (strain PAO1) DNA.
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
Pseudomonas.
REFERENCE 1 (sites)
AUTHORS Zimmermann, A., Reimann, C., Gallmand, M. and Haas, D.

TITLE Anaerobic growth and cyanide synthesis of *Pseudomonas aeruginosa* depend on *anr*, a regulatory gene homologous with *fnr* of *Escherichia coli*

JOURNAL Mol. Microbiol. 5 (6), 1483-1490 (1991)

MEDLINE 92157874

PUBMED 1787798

REFERENCE 2 (sites)

AUTHORS Savioz, A., Zimmermann, A. and Haas, D.

JOURNAL RpoN-independent promoters having a conserved GG-N10-GC motif in *Pseudomonas aeruginosa*

REFERENCE Unpublished (1992)

3 (bases 1 to 3761)

AUTHORS Zimmermann, A.

JOURNAL Anaerobe regulation des arginine-déaminase-opérons von *Pseudomonas aeruginosa* durch das protein *ANR*

Thesis (1992)

FEATURES location/Qualifiers

1..3761

/organism="Pseudomonas aeruginosa"

/strain="PA01"

/db_xref="taxon:287"

22..504

/note="ORF A; putative"

/number=1

/codon_start=1

/transl_table=11

/protein_id="AA25710.1"

/db_xref="GI:151020"

/translation="MDLIRWDADLIRYDLSGPRYSPTAVAFHEGIGEPDNLAL RDSRKAGHPLSLVHIFPCANICCYCACNKVITTKDGRSNAPYLARLVREIYSRHLS RAQVYQDLHFGGGTPTFLSPQDLRELSQLRHLNLDDSGDGLIDPREDWSTK"

CDS 564..1061

/note="ORF B; putative"

/number=2

/codon_start=1

/transl_table=11

/protein_id="AA25711.1"

/db_xref="GI:151021"

/translation="MPQRRINADDLPSPOKLEMLQRTTEFQLAAGYRIGMDHFLP DDELASQEDGTLOHFGYTHGHDIVGSAISQIDLSQNSDDINDQTSID NGQIAIRRLGHCHMSDDRRRAVYQQLICHELAFEDIEFEFGIDFSYFALMFDLER SPPTA"

complement(1235..1702)

/note="ORF X; putative"

/citation=[2]

/number=3

/codon_start=1

/transl_table=11

/protein_id="AA25712.1"

/db_xref="GI:151022"

/translation="MLTSGMTTVOVPPSSRAAADAARATGTAEQADPPAGTDAAD OHVLSQAGRELAATMGVPEEDEDIPQELRPMVMIRLKKIEKLELEAARS SPQAKAPARPELQKELQUNSLQTLATAMASAIKEMGISDPAIIKRWGSR"

1948..2682

/gene="anr"

1948..2682

/gene="anr"

/function="anaerobic regulatory protein analog to *FNr* of *E. coli*"

/citation=[1]

/citation=[2]

/number=4

/codon_start=1

/evidence-experimental

/transl_table=11

/product="ANR"

/protein_id="AA25713.1"

/db_xref="GI:151023"

/translation="MAETIKVRALPOAHCKDCSLAPLCLPLSTVEDMDSLDITVRKG RPLKGEFLERQDPPGSAVAVSGALKTFSITDAGEOITGFHLPSELVGLSGMDRE TVPVSQALETTSVCEIPERLDELSEQLPOLRQRLMSRLRDDQCMILLSKTQ ADRIRATPLVNLASRRARGFSAQOFRILANSRNEIGNYIGLAVEYTSRVTRFOQNGL

CDS 15AEGKEVHLIDSTIELCALAGGLEC"

2768..3316

/note="analog to the apt gene of *E. coli*; ORF C; putative"

/number=5

/codon_start=1

/transl_table=11

/protein_id="AA25714.1"

/db_xref="GI:151024"

/translation="MIFDEFTLKSQIRAVDPFPKPGVFRDITPLFQSPRALRMVDS FVQRYEADPSHIGAMDARGFLIGSNAVAVLANKLVFERKQKLPADVLAEGYQTEYG EARLEYHADSLCEGDSVLIJEDDLIATNGTLLAASIVRRIGARVFAAATIDPELGG SRILODAGISTESLTPALDER"

3356..>3361

/note="ORF D; putative"

/number=6

/codon_start=1

/transl_table=11

/protein_id="AA25715.1"

/db_xref="GI:151025"

/translation="M"

BASE COUNT 669 a 1337 c 1060 g 695 t

ORIGIN

Query Match 100.0%; Score 492; DB 1; Length 3761;

Best Local Similarity 100.0%; Pred. No. 1.8e-73;

Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTGCGGATCATCTTCACATCGGCGGCACTCTGCGGATATTCCTCGCTCTCTCT

DB 1450 ACTGCGGATCATCTTCACATCGGCGGCACTCTGCGGATATTCCTCGCTCTCTCT

QY 61 CCACCGGACCCCATGGTACGGGCAAGCTCGGCGGCTCTGGGAAAGCTGTACATGCT

DB 1510 CCACCGGACCCCATGGTACGGGCAAGCTCGGCGGCTCTGGGAAAGCTGTACATGCT

QY 121 GATCGCGGCGTGGTGGCGGCGGCGGCTTCCGCTGCTCGGCGGCTGGTCCGTG

DB 1570 GATCGCGGCGTGGTGGCGGCGGCGGCTTCCGCTGCTCGGCGGCTGGTCCGTG

QY 181 CGGCGTGGCGGCGGCGGCGGCGGCGGCTGAGGCGGCGGCTGGGTTGATCCAGCGAC

DB 1630 CGGCGTGGCGGCGGCGGCGGCGGCGGCTGAGGCGGCGGCTGGGTTGATCCAGCGAC

QY 241 TGAGGGTCAACATTCACATCGGCGGCAAAATGAAATTCCTTCATTCGATCGGCCAC

DB 1690 TGAGGGTCAACATTCACATCGGCGGCAAAATGAAATTCCTTCATTCGATCGGCCAC

QY 301 GCGTCGCACTTACAGCCCTTTTCTGTCGCGGCTTGACAGGTCGACAGTATGCGCA

DB 1750 GCGTCGCACTTACAGCCCTTTTCTGTCGCGGCTTGACAGGTCGACAGTATGCGCA

QY 361 GTTTTGGAGCAGTCACTGATTTGGAAGCGCATCGGCGGCTGAGAAATGCTGCTGCC

DB 1810 GTTTTGGAGCAGTCACTGATTTGGAAGCGCATCGGCGGCTGAGAAATGCTGCTGCC

QY 421 AGACCTATGCTGACAGCCGCGGCTGCTGATTCCTTCCTGCTGCTGCTTAA

DB 1870 AGACCTATGCTGACAGCCGCGGCTGCTGATTCCTTCCTGCTGCTGCTTAA

QY 481 CCTAGCAAGGAC 492

DB 1930 CCTAGCAAGGAC 1941

RESULT 3

AE004582/c 11627 bp DNA linear BCT 30-AUG-2000

LOCUS Pseudomonas aeruginosa PA01, section 143 of 529 of the complete genome.

ACCESSION AE004582 AE004091

VERSION AE004582.1 GI:9947492

KEYWORDS Pseudomonas aeruginosa.

SOURCE

ORGANISM Pseudomonas aeruginosa
Bacteria: Proteobacteria: gamma subdivision: Pseudomonadaceae;
Pseudomonas.
REFERENCE 1 (bases 1 to 11627)
AUTHORS Stover,C.K., Pham,X.O., Erwin,A.L., Mizoguchi,S.D., Warrenner,P., Hickey,M.J., Brinkman,F.S., Hufnagle,W.O., Kowalik,D.J., Lagrou,M., Garber,R.L., Goltry,L., Tolentino,E., Westbrook-Wadman,S., Iuan,Y., Brody,L.L., Coulter,S.N., Folger,K.R., Kas,A., Lardig,K., Lim,R., Smith,K., Spencer,D., Wong,G.K., Wu,Z. and Paulsen,I.T.
TITLE Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pathogen
JOURNAL Nature 406 (6799), 959-964 (2000)
MEDLINE 20437337
PUBMED 10984043
2 (bases 1 to 11627)
REFERENCE Stover,C.K., Pham,X.O., Erwin,A.L., Mizoguchi,S.D., Warrenner,P., Hickey,M.J., Brinkman,F.S., Hufnagle,W.O., Kowalik,D.J., Lagrou,M., Garber,R.L., Goltry,L., Tolentino,E., Westbrook-Wadman,S., Iuan,Y., Folger,K.R., Kas,A., Lardig,K., Lim,R.M., Smlth,K.A., Spencer,D.H., Wong,G.K., S., Wu,Z., Paulsen,I.T., Reizer,J., Saier,M.H., Hancock,R.E.W., Lory,S. and Olson,M.V.
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (16-MAY-2000) Department of Medicine and Genetics, University of Washington Genome Center, University Of Washington, Box 352145, Seattle, WA 98195, USA
FEATURES
source 1. 11627
/organism="Pseudomonas aeruginosa"
/strain="PAO1"
/db_xref="taxon:287"
70. .666
/gene="recr"
/note="PA1534"
70. .666
/gene="recr"
/codon_start=1
/transl_table=11
/product="recombination protein Recr"
/protein_id="AAG04923.1"
/db_xref="GI:9947493"
/translation="MSFPLIRLRIELESRLIPGVGKSNARMALMLERDRSGGLKLA QALTAAMEGVGHRCQCRITSEELCPCCADPRDLSLCYGPILDVAVVDTGGRG YFVLKGLHSPIDGSPPEAIGTPELEARRDRGAFSEVITATNPTVEGETAHTIAOLLA GRGLTSLRIAGVPLGSELVLDGTTAHLAAGRPI5"
789. .1937
/gene="PA1535"
789. .1937
/gene="PA1535"
/codon_start=1
/transl_table=11
/product="probable acyl-CoA dehydrogenase"
/protein_id="AAG04924.1"
/db_xref="GI:9947494"
/translation="MTDROVYFDESHOILRDSVRFVREVLPIIDMEAEPEFREL YLKAGACILIGTPEATGSGCEBDLFAKVAASEELMRKSGGLVAGLSIDIGLPV VKMAPEVREVRVPAVLNRGKELMALATVEPSGSDVANLKTAVNDGDHVRSGSKTF ITSCVRADYIVVAATGGEFGAGISLILVEKGTAGFSVGRKLKMGWMASTPAELFFD DCRYPAEVLIGVENAGFACIMANFOSEBRLALAVAMNTAOLALESLRMAREBEAFGK PIKGFQVLRHLAEATOLEVRSRETTYROAKMAAGSKVIEKISMAKNFADVDRLT YDAVOVLCGNGYMERSELYRLYRDRRIITISIGSGSEINETITGKMGJ"
complement(1950. .2489)
/gene="PA1536"
complement(1950. .2489)
/gene="PA1536"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAG04925.1"
/db_xref="GI:9947495"
/translation="MNPPLPHFRYHPEPLASGSIASATTCOCCKARGVYVYTGSPYS RHELPGSLCPMCIADGSAARAEASFSDYPLIDGVAADIIVTEVCERTGYSWOO ERMVLCEDACAFRGDAGREITGOLGAGLGAORADRFAMPAITWRLVADATTPGGRPA

gene TYRFDCLHCGAHYDLDT"
complement(2563. .3450)
/gene="PA1537"
CDS complement(2563. .3450)
/gene="PA1537"
/codon_start=1
/transl_table=11
/product="probable short-chain dehydrogenase"
/protein_id="AAG04926.1"
/db_xref="GI:9947496"
/translation="MKSFEKNVVAITGAGSGIGRALAVEIGROGCHLALADVNAALE ETROLASSGVAVSTAVVDVADREVOVMADKASERHVNILTFNNAGVAVNAGEGS DISEWIMINIFWGVNGTAKPLPHLASGHHVNVSVFGEAPQSGMAYNTATKY AVRGFTESLRQELMEDSDGSVCVHPGIGITNTIAKTARMNESMAKVTGAQPDAREQ FNDQLRTTPPEKAQVILRGVORDSRRLIGTDAHAIDVMLRLAPVLVQRIVTASMR LAAFAFRKSPGAREASE"
complement(3447. .5030)
/gene="PA1538"
complement(3447. .5030)
/gene="PA1538"
/codon_start=1
/transl_table=11
/product="probable flavin-containing monooxygenase"
/protein_id="AAG04927.1"
/db_xref="GI:9947497"
/translation="MYTPANNHNSLAMPSTPTPAARHCKVAIIGTGFSGIGNAIRL ROEGEDDLFIEKADAGVGTWRVNNYPCGACDVOSHVYSFSEANPEMTRMFARPEI RAYLECKEYKRLQKTLNTEIGKLAMDROSLMHLHDAOGNHTYTAAYVSGMGLS TPAYLRDGLDENFGKVFHSGOQMDHVDLKKRAVAVIGTGAATQFPEIPIVAALD LYQRTPEVILRPDPAISETERRRRRPLVOKLMRGSLYSLBGRVUGTFPAPQVK LVORLAIRHKOIKDPELRRKVTVDYTGKRIIMSHNYPALAANSTVITTEIRA VTANGIVDNGEREVEDAIIFGTGTADPPIRGVFGDRDLDSKQPEAVKGT TTAGFPNLFLMGPNTGGLGSHNSMYMVISQIAYVADALKMLKRLRLSLSEKAPQER YNEYLQRLKIDRSVMSVSGCKSWYLHPVSGRNCTIIMPFTWFRPALTREQDASAVHLTT TPLAALSNEARQACGVPA"
5144. .5992
/gene="PA1539"
5144. .5992
/gene="PA1539"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAG04928.1"
/db_xref="GI:9947498"
/translation="MSDMNETKTRDVAAPGRKALILEALRLGSSSRSLGATGLRELAR EAGLNPTFYRHFRIDIDLGLTMRIDISTOLRQPLRLRRAATPAAGARAQTTPTG LDIERGRVRCETVRLFDVFNESNPAATIVGVRELHGAAPVLRGLARQIMPEFADMS EDLIEFLTPAVSEATIRLSLISIGRQFLQSLDYIGQPERRETCAGEEOIILLFT GAVYQMMGVSLDASGQLEQAHAGQ000QAEPAPQALAEQRPAGHDGDEAEAA PDRIGE"
complement(5846. .6175)
/gene="PA1540"
complement(5846. .6175)
/gene="PA1540"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAG04929.1"
/db_xref="GI:9947499"
/translation="ALTIINWTFPMALGAIALLEVYANILLLKYSDFRRRGIGINSIIC VMAFTALAQAVKDIELSAYAIWGFILATVAMGALFGORLAMRWGLDLLLLLAG MSLKLA"
complement(6169. .6537)
/gene="PA1541"
complement(6169. .6537)
/gene="PA1541"
/codon_start=1
/transl_table=11
/product="probable drug efflux transporter"
/protein_id="AAG04930.1"
/db_xref="GI:9947500"
/translation="MRSWYIILLIAGAVYIGTTSKMLAATHTAPVAGMLIAMGIGLSY FFLAIVARVPVGAVALMEGIGIVLITVASVAMGESIGLYKAVGLGVMTAGILLIK

	gene	SCRNAAGTPTAOSGEAVTC" complement(6946. .7782) /gene="PA1542"
	CDS	/complement(6946. .7782) /gene="PA1542" /codon_start=1 /transl_table=11 /product="hypothetical protein" /protein_id="AA04931.1" /db_xref="GI :9947501" /translation="MTATGTASPYRANFPVRRMDFSESPKRYWMDGCEPFTOPMNIL SLPEYGEKFEPVDRAVRDRLEDDPOLKDVSAITGOEAMSKENHAYINIEAEHGVD LECELEIKVILENKSQVTTRKHOLATCALEHTTAAMAOLKREDTLTOMOSPKNAT KLMMHAHIEBHEKNVAIVDYQKVYGGTFETRVLMWLTTLMFEVIGCFOLDLRKCG QLFWMRSWKHGLTLFLGRGFETRLLLPLYDYYRPQHPSDHDTAKLEKYMKDLAR AD"
	gene	complement(8022. .8570) /gene="apt" /note="PA1543"
	CDS	complement(8022. .8570) /gene="apt" /codon_start=1 /transl_table=11 /product="adenine phosphoribosyltransferase" /protein_id="AAC04932.1" /db_xref="GI :9947502" /translation="MFDEFTLKSQLRAVPDPFKGVVFRIPIPLFOSPRALRTAVDS FYVGVIADSFHSICDMARGFLISAVAVALNRVLFRKGKLPADVLAEGYOTERYGA FAFLVEHADSHICEGDSVLIIPDDLATGTLIAASTLVRLGARVFEMAAILIDLELGG STRQQDGISTPESSTAFALDER" complement(8656. .9390) /gene="anr" /note="PA1544"
	CDS	complement(8656. .9390) /gene="anr" /codon_start=1 /transl_table=11 /product="transcriptional regulator Anr" /protein_id="AAC04933.1" /db_xref="GI :9947503" /translation="MAETIKYRALPOAHCKDCSLAPICLPILTVEEDMSIDEIYKRKG RPLKKGEFLPROGDPPRGSSVRAVRSGALKTTSITDAGEOITGHLSBELVGLSMDIE TYVPSAOALETTSVCLEIFERDELDSLQPLQRKQLMRSLREIRDOQMMLLSKTTE ADEIAETFLVNLSAFRAPRFESAQQFPLAMSRLNEIGNYLGLAVETVSRLFQOONGL ISAEGKEVIIIDSLTEICALAGGLEG" 9636. .10103 /gene="PA1545"
	Query Match	99.7%; Score 490.4; DB 1; Length 11627;
	Best Local Similarity	99.8%; Pred. No. 2.9e-73;
	Matches 491; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
OY	1	ACTGCGGATATCATCTTCACATCGGCCCCCACTCCTGGGGGATATCCCTGCCTCTCT
Db	9888	ACTCGGGATATCATCTTCACATCGGCCCCCACTCCTGGGGGATATCCCTGCCTCTCT
OY	61	CCACCGGACCCCCCATGTAGTACGGGACACCTGCGGCTCGCTTGGAAGAAGCTGTACATGCT
Db	9828	CCACCGGACCCCCCATGTAGTACGGGACACCTGCGGCTCGCTTGGAAGAAGCTGTACATGCT
OY	121	GATCGGCGGGCGTGGTGGCCGGCGGGCGGGGCTTCCCGCTGCTCGGCGGTGCCGGTCCGGTG
Db	9768	GATCGGCGGGCGTGGTGGCCGGCGGGCGGGGCTTCCCGCTGCTCGGCGGTGCCGGTCCGGTG
OY	181	CGGCGTTCGGCGTCCGGCGGGCGGGCGGCATGAGAGGGGACACTGGGTGATCCAGCCAC
Db	9708	CGGCGTTCGGCGTCCGGCGGGCGGGCGGCATGAGAGGGGACACTGGGTGATCCAGCCAC
OY	241	TGAGGATCAAGATTCCAATCACTCCGGGAAAAAATGGAATTCTTCATTGGATCGGCCAC
Db	9648	TGAGGATCAAGATTCCAATCACTCCGGGAAAAAATGGAATTCTTCATTGGATCGGCCAC
OY	301	GGCTCGCAACTTGAGCCCCCTTTGTCGCCCTTGACAGAGGTCCGACAGGTAGTCCGA

Db	9588	GGGTGCGCAACTGTAGACCCCTTTTCGTCGCCCCCTTGACAGGGTGTCCAGAGTACTGCCA	9529
OY	361	GTGTGTTGACGCACTCACTGATTTGGAAGCCGATCGGCTGTGAGAAATGGTCGTTGCC	420
Db	9528	GTTGTTTACGGAATCACTGATTTGGAAGCCGATCGGCTGTGAGAAATGGTCGTTGCC	9469
OY	421	AGACCTATGCGTGGCACCAGCCGATCGGCTGTGAGAAATGGTCGTTGCC	480
Db	9468	AGACCTATGCGTGGCACCAGCCGATCGGCTGTGAGAAATGGTCGTTGCC	9409
OY	481	CCTACGACGAC	492
Db	9408	CCTACGACGAC	9397

LOCUS	PAHEMN	2635 bp	DNA	linear	BCT 29-AUG-1996
DEFINITION	P. aeruginosa hemN gene.				
ACCESSION	X97981				
VERSION	X97981.1	St.1515302			
KEYWORDS	ANR-binding motif; coproporphyrinogen III dehydrogenase; hemN gene; ORFX.				
SOURCE	Pseudomonas aeruginosa.				
ORGANISM	Pseudomonas aeruginosa				
REFERENCE	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae; Pseudomonas.				
AUTHORS	1 (bases 1 to 2635)				
TITLE	Hungerer,C., Troup,B., Raemling,U. and Jahn,D.				
JOURNAL	Cloning and regulation of the Pseudomonas aeruginosa hemN gene				
REFERENCE	encoding an oxygen-independent coproporphyrinogen III dehydrogenase				
AUTHORS	2 (bases 1 to 2635)				
JOURNAL	Jahn,D.				
REFERENCE	Direct Submission				
AUTHORS	Submitted (20-MAY-1996) D. Jahn, Laboratorium fuer Mikrobiologie,				
JOURNAL	Fachbereich Biologie, Philipps-Universitaet Marburg,				
REFERENCE	Karl-von-Frisch-Str., D-35032 Marburg, FRG				
AUTHORS	location/Qualifiers				
JOURNAL	1. 2635				
REFERENCE	/organism="Pseudomonas aeruginosa"				
AUTHORS	/strain="PAOI"				
JOURNAL	/db_xref="taxon:287"				
REFERENCE	632. 645				
AUTHORS	/bound_moiety="ANR"				
JOURNAL	709. 2091				
REFERENCE	/gene="hemN"				
AUTHORS	709. 2091				
JOURNAL	/gene="hemN"				
REFERENCE	/EC_number="1.3.3.3"				
AUTHORS	/codon_start=1				
JOURNAL	/transl_table=11				
REFERENCE	/product="oxygen-independent coproporphyrinogen III				
AUTHORS	dehydrogenase"				
JOURNAL	/protein_id="CA66617.1"				
REFERENCE	/db_xref="GI:1515308"				
AUTHORS	/db_xref="SWISS-PROT:p77915"				
JOURNAL	/translation="MDTIRKADLLIRYDISGRYISYTPANVFHEIGIEFDQIALH				
REFERENCE	RDSKAGHPISLVIHIFPCANIYYCACNKVITIDKRSAPYIARLVLEIYSRLHS				
AUTHORS	RAQVVEQLFGCGTPTPLSPGOLRELSQITHTNLDDSGDYGIEDREADSTM				
JOURNAL	GLAELGFNRYSLGVQEDFDMVKAVNRMQPPEETRIVEAKRTLOYRSINLDLYGL				
REFERENCE	PKQKDFPARTVDEVIALQPRMSVFNVAHLPERFMQRTINADLPSPQKLEMLOR				
AUTHORS	TTREOLAAGRYRISMDHFAIPDDLSAENOEGTIGQRNFGCTTHGKDLVGLGYSATS				
JOURNAL	QTEGIFSONSSINDIYQTSIDNGQLAIRGLHCSDDVRRAVVIQILCHELAIEDIT				
REFERENCE	ETEGEIDRSYFAELWDLERFADGLIRPDAKGIDITSSGRLVRSICMLFRYIPSP				
AUTHORS	LNRQFRSVIT"				
JOURNAL	complement(2140..2607)				
REFERENCE	/note="ORFX"				
AUTHORS	/codon_start=1				
JOURNAL	/transl_table=11				
REFERENCE	/protein_id="CA66618.1"				
AUTHORS	/db_xref="GI:1515309"				

BASE COUNT	ORIGIN
421 a	971 c 779 g 403 t 1 others

Query Match	281;	Score	281;	DB	1;	Length	2635;
Best Local Similarity	100.00;	Pred. No.	5.6e-38;				
Matches	281;	Conservative	0;	Mismatches	0;	Indels	0;
						Gaps	0

QY	1	ACTCGCGGATCATCTTACACATCGGCGCGCAACACTCTGGGGATATCTGTCTCTCTCT	60
Db	2355	ACTCGCGGATCATCTTACACATCGGCGCGCAACACTCTGGGGATATCTGTCTCTCTCT	2414

Db	Seq	Length
61	CCACGGGACCCCATGGTAGCGGACAGCTGGGCCCCCTGGGAAACCTGTACATGCT	120
2415	CCACGGGACCCCATGGTAGCGGACAGCTCGGCCCCCTGGGAAACCTGTACATGCT	2474

	gq	121	GATCGGCGGGCGTGGGTGCACGGCCGGGCTTCCTCAGCTCGTCGACGGCGGATGCCGGTG	180
	Db	2475	GATCGGCGGGCGTGGGTGCACGGCCGGGCTTCCTCAGCTCGTCGACGGCGGATGCCGGTG	2534

Db	CGGCTTGGCGTCGGCGGCCGCAGTAGAGGCGGCACCTGGGTGGTGATCCAGCCAC	2594
QY	CGGCTTTGGCGTCGGCGGCCGCAGTAGAGGCGGCACCTGGGTGGTGATCCAGCCAC	240
181	CGGCTTTGGCGTCGGCGGCCGCAGTAGAGGCGGCACCTGGGTGGTGATCCAGCCAC	240

Db	2595	TGAGGGTCACATTC	2635
0y	241	TGAGGGTCACATTC	281

RESULT 5	PAANR	LOCUS	PAAND	900 bp	DNA	1400000	PCMC 14-ATG-1000

DEFINITION
P.aeruginosa anr gene for a transcriptional activator protein.
ACCESSION
X58405
VERSION
X58405.1 GI:45377
FEATURES
KEYWORDS
anr gene, transcriptional activator

SOURCE ORGANISM
pseudomonas aeruginosa.
pseudomonas aeruginosa
Bacteria; Proteobacteria; gamma subdivision; pseudomonadaceae; pseudomonas

REFERENCE
1 (pages 1 to 900)
SAVERS, R.G.
Direct Submission
Submitted (11-mar-1991) P C Savers
Title
Abstract
Journal

REFERENCE
2 (Pages 1 to 900)
UNIVERSITÄT MÜNCHEN, MARIA WARD STR 1a, 8000 MÜNCHEN 19, GERMANY
AUTHORS
SAVERS, R.G.
TITLE
IDENTIFICATION AND MOLECULAR CHARACTERIZATION OF A TRANSCRIPTIONAL

JOURNAL
OF
MOLECULAR
BIOLOGY
157, 1469-1481 (1991)

FEATURES	COMMENT	PUBMED
transcriptional activator of anaerobic gene expression which shows a high degree of amino acid similarity to FNR protein of E.coli.		1/8/11/9/

```
source
1. .900
/organism="Pseudomonas aeruginosa"
/strain="PA01"
/db xref="taxon.387"
```

```
gene          66. 831
              /gene="anr"
prlm_transcript 66. >831
              /gene="anr"
```

```

CDS
    57..631
    /gene="anr"
    /codon_start=1
    /transl_table=11

```

```

/product="transcriptional activator protein"
/protein_id="CAA41305.1"

```

```

/db.xref="GI:45278"
/db.xref="SWISS-PROT:P23926"
/translation="MAETIKVRAIPQAHKDCSLAPLCPLSTVEDMDSLDEIVKRG
RPLKGEFLPROGPGSVAFNRGSLAKTESINDAGEEQITGFGLPELIGLSOMDE
TYPSSAQLETTSVCELPFERLDLSLQPLQRLQRLMLRSRLRDDQAMLLSKTT
ADEISIAFLNLSARFRARGFSAQOFLNARINEIGYGLAVEYVSRYVTRFQONL
```

BASE COUNT ORIGIN	165 a	310 c	261 g	164 t
----------------------	-------	-------	-------	-------

Best Local Similarity 100.0%; Pred. No. 8 2e-05;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1 TCAGAAATGCTGCTTGCACAGACTATGCTGGCACCCCGCATCGCGGCTGCGTTACCTTA 60

Db 61 CTCCTGTGTCCTTACCTAGCAAGAC 90

RESULT 6	
AC087547/c	
LOCUS	
AC087547	
129420 bp	
DNA	
linear	
PLN 02-MAR-2001	

ACCESSION	ncbe0040111, complete sequence.
VERSION	AC087547
	AC087547.2
	GI:13184902

SOURCE ORGANISM
Oryza sativa (japonica cultivar-group).
Oryza sativa (japonica cultivar-group).
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE
AUTHORS

TITLE
 Rice Chromosome 10
 JOURNAL
 Unpublished
 REFERENCE
 2 (bases 1 to 129420)
 2 (bases 1 to 129420)

TITLE
Nevill-Manning, C. and Messing, J.
Direct Submission
Submitted (08-JAN-2001) The Plant Genome Initiative at Rutgers -
The State University of New Jersey, New Brunswick, NJ 08902

REMARK	REFERENCE
Piscataway, NJ 08873 Chromosome 10 3 (bases 1 to 129420)	

TITLE
Messing, J.
Direct Submission
Submitted (02-MAR-2001) The Plant Genome Initiative at Rutgers -
The State University of New Jersey
JOURNAL
Published (02-MAR-2001) The Plant Genome Initiative at Rutgers -
The State University of New Jersey

REMARK	Piscataway, NJ 08873
COMMENT	Chromosome 10
REMARK	On Mar 2, 2001 this sequence version replaced gi:12044846.

```
source  
1..129420  
/organism="Oryza sativa (japonica cultivar-group)"  
/sub-species="japonica"  
/db-species="japonica"  
/db-species="japonica"
```

```

/chromosome="10"
/clone="nbe0040111"
BASE COUNT      35961 a      28213 c      30048 g      35198 t
ORIGIN

```

Query Match	10.5%	Score 51.8;	DB 8;	Length 129420;
Best Local Similarity	52.6%	Pred. No. 10;		
Matches 113	Conservative	0	Mismatches 103	Indels 0
			Cuts 0	

oy 14 CTTACCATCGGCCGAACCTCCTGGGGATATCTCTCTCTCTCTCCACCGGCACCCC 73

[illegible]

OY	14	CTTACCAATCGGGCCGCAACTCCTGGGGGAATATCTTCGTCTCTCTCTCACCGGCACC	73
D5	35415	CTCGATCATGCCGCGAGACGTTTGAAACACCGCGCGCGCCAGCAGCAGCAGCATT	35356
OY	74	CATGATTAACGGCGACATCGCGCCCTGCGGGAAGAGCTATATCTGATCGGGGCGTC	133
D5	35355	CACGCGCTGTCTCCAGACACAGCAGCAGCTTTCGCGCTCTCTCTCCGCGGCGGGGG	35296
OY	134	GGTGCGGCGCGCGGATCTTCGCGCTGTCTCGGCGGTGCGGTCTGCGGCTTGGCGTC	193
D8	35295	CGGCGCAGCGCGGTGCTGTGTCGCGCTCTCCGCGGCGAGGGGGCGCTTTCCTTGTGTC	35236
OY	194	CGCGGCGCGCGCGCATGAGGCGCGCACCCTGGCTG	228
D8	35235	CTCGCGCGCGCGGCTTGTGAGGCGCATGAGTGTGG	35201
RESULT 8			
AP003771/c			
LOCUS	AP003771	175644 bp	DNA linear HTG 21-MAR-2002
DEFINITION	Oryza sativa (japonica cultivar-group) chromosome 6 clone P0622F03,		
ACCESSION	** SEQUENCING IN PROGRESS ***, in ordered pieces.		
VERSION	AP003771		
KEYWORDS	AP003771.1 GI:14517645		
SOURCE	HTG; PHASE2.		
ORGANISM	Oryza sativa (japonica cultivar-group) (cultivar:Nipponbare) DNA,		
REFERENCE	clone:P0622F03.		
AUTHORS	Oryza sativa (japonica cultivar-group)		
TITLE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
JOURNAL	Spermatophytes; Magnoliophyta; Liliopsida; Poales; Poaceae;		
AUTHORS	Ehretidae; Oryzae; Oryza.		
REFERENCE	1 Sasaki,T., Matsumoto,T. and Yamamoto,K.		
TITLE	Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, PAC		
JOURNAL	clone:P0622F03		
AUTHORS	Published Only in Database (2001)		
REFERENCE	2 (bases 1 to 175644)		
TITLE	Sasaki,T., Matsumoto,T. and Yamamoto,K.		
JOURNAL	Direct Submission		
COMMENT	Submitted (20-JUN-2001) Takui Sasaki, National Institute of Agrobiological Resources, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan (E-mail:tsasakigabr.afric.go.jp, URL:http://rgp.dna.afrc.go.jp/, Tel:81-298-38-7441, Fax:81-298-38-7468) NOTE: It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the places is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter. This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved. * NOTE: This is a 'working draft' sequence. * This sequence will be replaced * by the finished sequence as soon as it is available and * the accession number will be preserved. location/Qualifiers		
FEATURES	source		
BASE COUNT	46662 a 41378 c 40541 g 46713 t 350 others		
ORIGIN	1..175644		
	/organism="Oryza sativa (japonica cultivar-group)"		
	/cultivar="Nipponbare"		
	/db_xref="taxon:3994"		
	/chromosome="6"		
	/clone="P0622F03"		
Query Match	10.1%; Score 49.8; DB 2; Length 175644;		
Best Local Similarity	53.3%; Pred.No. 21;		
Matches 105; Conservative	0; Matches 92; Indels 0; Gaps 0;		
OY	2	CTCGGGATCATCTTACCATCGCGGCAATCCCTGCGGGAATATCTCTCTCTC	61
D5	85945	CCCCGGAAGCTCCGACACCCGCGCTCTCTTACACGCGCGCTCTGTCGACAA	85886

```

OY      62 CACCGGCAACCCCATGTAGACGGCAGCTGCCTGCCCTGGGAAGTGTACATGCG 121
DB      85885 GCCGGTGCCGTCCTTCCTCCGCCACCCTAGACTGTCCCTGCGACAGCTCTTTCCGTC 85826
OY      122 ATCGGCGGCGTCGCGTCCGCGCGCGGGGTCTTCCGCTGTCGGCGGTCCGCTCCGTC 181
DB      85825 ATTGCGCGTCGTGTGTGTGTGTGTGTCGTCGTCCTCCCTCCAAAGGGACCGGACAC 85766
OY      182 GCGCTTGGCGCTCCGCGG 198
DB      85765 GCCGAGGAGGTCCGCGG 85749

RESULT 9
AC090435/c
LOCUS    AC090435          68736 bp        DNA           linear       HTG 25-JUL-2002
DEFINITION Chlamydomonas reinhardtii clone cr-32m22, WORKING DRAFT SEQUENCE, 3
ordered pieces.
ACCESSION AC090435
VERSION   AC090435.25  GI:21747477
KEYWORDS  HTGS_PHASE2; HTGS_DRAFT.
SOURCE    Chlamydomonas reinhardtii.
ORGANISM  Chlamydomonas reinhardtii
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
Chlamydomonadales; Chlamydomonas.
REFERENCE
AUTHORS   Wu,H., Lin,S., Jia,H., Dutcher,S. and Roe,B.A.
TITLE     Chlamydomonas reinhardtii BAC Clone cr-32m22
JOURNAL   Unpublished
REFERENCE
AUTHORS   Wu,H., Lin,S., Jia,H., Dutcher,S. and Roe,B.A.
TITLE     Direct Submission
JOURNAL   Submitted (21-FEB-2001) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE
AUTHORS   Mu,H., Lin,S., Jia,H., Dutcher,S. and Roe,B.A.
TITLE     Direct Submission
JOURNAL   Submitted (25-JUL-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT   -- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 1364: contig of 1364 bp in length
* 1365 13464: gap of unknown length
* 13465 32124: contig of 18660 bp in length
* 32125 32224: gap of unknown length
* 32225 68736: contig of 36512 bp in length.
Location/Qualifiers
1..68736
/organism="Chlamydomonas reinhardtii"
/db_xref="taxon:3055"
/clone_lib="cr-32m22"
/clone_1ib="cr-32m22"
BASE COUNT 11440 a 23005 c 23233 g 10858 t 200 others
ORIGIN
Query Match 10.0%; Score 49.4; DB 2; Length 68736;
Best Local Similarity 57.4%; Pred. No. 28;
Matches 89; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
```

OY		49	CCTCCTCCATCCACCGGCAAGGGCATGTTAGCGGCAGTCGCGCCCTTGCTGGAAA	108
Dd	44925	CGCGGCCGCGCGCGCGCGCGCGCGGTGACGTGCAAGTCCCGGCTGCGCGGCGG	44866	
OY	109	GCTGACATGCTGATTGCGGCGGCGTGCATGCGCGCGCGCGCGCGGCTTCCTGCTGCGCGG	168	
Dd	44865	CTGCGCGCGCTGATGCGCGCGCGGTGAAGCGCGCGCGCTGCGGTGCGAGGTGCGCGCGCGG	44806	
OY	169	TGCCGCTCCGTCGCGCGCTTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	203	
Dd	44805	CGCGCGCGCGCTGCAAGCG	44771	
RESULT	10			
LOCUS	AC090436			
DEFINITION	Chlamydomonas reinhardtii clone cr-3hl, WORKING DRAFT SEQUENCE,			
ACCESSION	AC090436			
VERSION	AC090436.33			
KEYWORDS	HTGS, HTGS_PHASE2; HTGS_DRAFT.			
SOURCE	Chlamydomonas reinhardtii.			
ORGANISM	Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;			
REFERENCE	Chlamydomonadaceae; Chlamydomonas.			
AUTHORS	1 (bases 1 to 102657)			
JOURNAL	Jia,H., Lin,S., Wu,H., Dutcher,S. and Roe,B.A.			
REFERENCE	Chlamydomonas reinhardtii BAC Clone cr-3hl			
AUTHORS	Unpublished			
JOURNAL	2 (bases 1 to 102657)			
REFERENCE	Jia,H., Lin,S., Wu,H., Dutcher,S. and Roe,B.A.			
AUTHORS	Direct Submission			
JOURNAL	Submitted (21-FEB-2001) Department Of Chemistry And Biochemistry, The University of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA			
REFERENCE	3 (bases 1 to 102657)			
AUTHORS	Jia,H., Lin,S., Wu,H., Dutcher,S. and Roe,B.A.			
JOURNAL	Direct Submission Submitted (04-AUG-2002) Department Of Chemistry And Biochemistry, The University of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA			
COMMENT	On Aug 3, 2002 this sequence version replaced gl:21747478. ----- Genome Center Center: Department Of Chemistry And Biochemistry The University Of Oklahoma Center code:UOKNOR -----			
FEATURES	* NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submittor. This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved. 1 102657: contig of 102657 bp in length. Location/Qualifiers 1..102657 /organism="Chlamydomonas reinhardtii" /db_xref="taxon:3055" /clone="cr-3hl" /clone_1lb="cr-3hl"			
BASE COUNT	17795 a 32817 c 34503 g 17535 t			7 others
ORIGIN				
Query Match	10.0%; Score 49.4; DB 2; Length 102657; Best Local Similarity 57.4%; Pred. No. 26;			
Matches	89; Conservative 0; Mismatches 66; Indels 0; Gaps 0;			
OY	49	CCTCCTCCTCCTCCACCGGCAAGGGCATGTTAGCGGCAGTCGCGCCCTTGCTGGAAA	108	

ANQATPAPNNKCTVYKSGNGOVNANALAIYFTGOTKRTVYALDYLRNP
 LNSTITGTEGEPVNRHHRWAHQADSLPAPGSLAGPNSLQDPLAAAOIAAC
 APOKCYVSHIAYSLVIMNSALVIMASAMPBRADARA"

BASE COUNT 339 a 949 c 885 g 277 t

Query Match

Best Local Similarity 51.6%; Score 48.2; DB 1; Length 2450;
 Matches 110; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 26 CCGCACTCTGCGGATATCTGCTCTCTCCACCGGACCCCAATGATGCGG 85
 DB 1800 CCGGTAAGACGGGAGCCGCTGACGCGGCGGATACCGGCTCGCTGCGGAA 1741
 QY 86 CAGCTCGCGCCCTGCTGCGGAAAGCTGATACATGCTGATGCGGCGGCTGCGGCGG 145
 DB 1740 CAGCTCGCGGCGGCGGCGGAGTACGAGTACGAGGAGGCGGCGGCGGCGGCGG 1681
 QY 146 CCGGCTTCTCCGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 205
 DB 1680 CCGGCTTCTCCGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1621
 QY 206 CGATGAGCGCGGACCTGCGGCTGATCCAGCC 238
 DB 1620 GCGCGCGGCGGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1588

RESULT 13
 AF355797 2668 bp mRNA linear PRI 16-JUL-2002

LOCUS Homo sapiens BTBD2 protein mRNA, complete cds.
 DEFINITION AF355797
 ACCESSION AF355797.1 GI:13430407

VERSION AF355797.1
 KEYWORDS human.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 2668)
 Xu, L., Yang, L., Hashimoto, K., Anderson, M., Kohlhaagen, G., Pommer, Y.,
 and D'Arpa, P.
 TITLE Characterization of BTBD1 and BTBD2, two similar
 BTB-domain-containing Kelch-like proteins that interact with
 Topoisomerase I

JOURNAL BMC Genomics 3 (1), 1 (2002)
 PUBMED 11818025
 REFERENCE 2 (bases 1 to 2668)
 Xu, L., Yang, L., and D'Arpa, P.
 TITLE Direct Submission

JOURNAL Submitted (02-MAR-2001) Biochemistry & Molecular Biology, Uniformed
 Services University of the Health Sciences, 4301 Jones Bridge Road,
 Bethesda, MD 20814-4799, USA

FEATURES
 source 1..2668
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

CDS
 1..1578
 /note="BTB-containing Kelch-like protein"
 /codon_start=1
 /product="BTBD2 protein"
 /protein_id="AK25826.1"
 /db_xref="GI:13430408"

translation="MAAGSGGASCPGPGVGTGSGSPGSANANAATPAPGNA
 AAAAAAAGPTPPAPGPTDAAGAEAEAGAGAAALQREAYNNQAKPT
 VOERAFPFNNEVLCDVFLVKGSLSORIPARHPLVAGSAVDFAMFGCAATSTE
 TELPEVPAALFALIKPLYSDEVOIGPEVTTLTAKKAYPALEACHVEIKKTLR
 ADNAEMILTOARLDEPOLASICLENIDKNADATAGFTDIDITLVAVIEROTLG
 IREYRLNNAVWRMSAEQOROLQVTPERRKVLKALGLINFLPLTIEEPAGPAQS
 GLVDREVSLVLEHTVNPKPVEFIDRRPCLRKESINFGQVESWGSCTSDR
 IFLVYKRIIFVVGGLGSHGPTDQVNIQIIRHDSNTVLGQNDTGFSCDSASTFR
 VMFEKPEVLVNTYACATLKGPSDHSHTKGLRKVTHTESPPTGAKCTETCYAAGNN
 NGTSEVDOIEVEIYTT"

BASE COUNT 487 a 932 c 820 g 429 t

ORIGIN

Query Match 9.8%; Score 48; DB 9; Length 2668;
 Best Local Similarity 53.8%; Pred. No. 79;
 Matches 99; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 23 CCGCGGCACTCTGCGGATATCTGCTCTCTCCACCGGACCCCAATGATGCGG 82
 DB 264 CACG 205
 QY 83 GCGCAGCTGCG 142
 DB 204 GTCTGCTCCG 145
 QY 143 GCGCGGCTTCTCCGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 202
 DB 144 GCGCGGCGCGCGCGCGCGCGCGCGGCTTCCGCGGCGCGCGCGCGCGCGCG 85
 QY 203 GCGC 206
 DB 84 GCGC 81

RESULT 14
 AC004678 38939 bp DNA linear PRI 12-MAY-1998

LOCUS Homo sapiens chromosome 19, cosmid R34094, complete sequence.
 DEFINITION AC004678
 ACCESSION AC004678
 VERSION AC004678.1 GI:3128156
 KEYWORDS HTG.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 38939)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 AUTHORS Lamerdin, J.E., McCreedy, P.M., Skovronski, E., Adamson, A.W.,
 Burkhardt-Schultz, K., Gordon, L., Kyle, A., Ramirez, M., Stillwagen, S.,
 Phan, H., Velasco, N., Barnes, J., Dangnanan, L., Poundstone, P.,
 Christensen, M., Georgescu, A., Avila, J., Liu, S., Attix, C.,
 Andreise, T., Trankheim, M., Amico-Keller, G., Coefield, J., Duarte, S.,
 Lucas, S., Bruce, R., Thomas, P., Quan, G., Krommiller, B., Arellano, A.,
 Montgomery, M., Ow, D., Nolan, M., Trong, S., Kobayashi, A., Olsen, A.S.,
 and Carrano, A.V.

TITLE Sequence analysis of a 3.5 Mb contig in human 19p13.3 containing a
 serine protease gene cluster
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 38939)
 Lamerdin, J.E.
 TITLE Direct Submission

JOURNAL Submitted (12-MAY-1998) Joint Genome Institute, Lawrence Livermore
 National Laboratory, 7000 East Ave., Livermore, CA 94551, USA
 COMMENT R34094 overlaps cosmid R27216 to the left and R20063 to the right.
 Additional chr 19 map and sequence information may be obtained at:
 http://www.bio.livnl.gov/hbip/genome/genome.html.

FEATURES
 source 1..38939
 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"

/chromosome="19"
 /map="19p13.3 between CDC34 and D19S342"
 /clone="R34094"
 /cell_line="5HL2-B"
 /clone_11b="L119NC03 R chromosome 19-specific cosmid
 library"

/note="Cosmid library constructed at LLNL from flow-sorted
 chromosomes from hybrid 5HL2-B, which carries chromosome
 19 as its only human chromosome."

misc_feature
 complement(82..204)
 /note="DGS similarity to T50998 yb71905.s1 Homo sapiens
 cDNA clone 76664.3, contains MER22 repetitive element;
 (123..1); 96% identity."

repeat_region 83..239

```
repeat_region /rpt_family="Alusq"
complement(240..5477)
repeat_region /rpt_family="Alusx"
complement(555..687)
repeat_region /rpt_family="FLAM.C"
complement(727..1019)
repeat_region /rpt_family="Alusx"
complement(1058..1165)
/rpt_family="GC-rich"
complement(join(<1155..1324,5492..5611,23445..235622))
CDS /note="Hypothetical human protein most similar to (277660)
F38H4.7 [Caenorhabditis elegans]"
/codon_start=3
/product="R34094_1"
/protein_id="AAC16070.1"
/db_xref="GI:3128157"
/translation="ASKPTVOERAPFLFNNVLCGVHFNKGLSSQRIAPRFLVAV
GSAYFDAMFNGMATTSRETEIPVEPAFLALKFLXSDVQIGPETWTLTTAKK
YAVPALAHCEVFEFLKKNLRADNAFMLITOVGA"
misc_feature complement(1155..1324)
/note="DPS similarity to PID1e255903 (277660) F38H4.7
[Caenorhabditis elegans]; (residues 253..309); 67%
identity. predicted exon, program: grail2exons.human_1.3,
frame: 2, quality: excellent, score: 100.000-DDS
similarity to overlapping ESTs:
(1324..1168) T50998 yb71905.s1 Homo sapiens cDNA clone
76664.3' contains MER22 repetitive element; (280..124);
96% identity.-(1324..1167) W62177 m87d04.r1 Soares mouse
embryo NbME13.5 14.5 Mus musculus cDNA clone 375367.5'
similar to WP:116H12.5 CE00510 KRUPPEL-LIKE ZINC FINGER
PROTEIN ; (134..311); 87% identity.-(1324..1168) T51086
yb71905.r1 Homo sapiens cDNA clone 76664.5'; (61..221);
98% identity.-Other EST matches: T51086"
repeat_region /rpt_family="Alusx"
2184..2483
/rpt_family="Alusx"
2497..2612
repeat_region /rpt_family="FLAM.C"
2628..2926
/rpt_family="AlusC"
complement(2937..3103)
repeat_region /rpt_family="LI"
complement(3114..3414)
repeat_region /rpt_family="Alusx"
complement(3428..3729)
repeat_region /rpt_family="Aluv"
complement(3785..4102)
repeat_region /rpt_family="Aluv"
complement(4107..4639)
/rpt_family="LI"
4659..4713
repeat_region /rpt_family="AlusP/q"
4714..4747
/rpt_family="(GAAA)n"
4750..5037
repeat_region /rpt_family="Alusx"
5049..5343
/rpt_family="Aluv"
complement(5492..5611)
/note="DPS similarity to PID1e255903 (277660) F38H4.7
[Caenorhabditis elegans]; (residues 209..252); 72%
identity. predicted exon, program: grail2exons.human_1.3,
frame: 2, quality: excellent, score: 95.000-DDS similarity
to overlapping ESTs:
(5555..5492) T50998 yb71905.s1 Homo sapiens cDNA clone
76664.3' contains MER22 repetitive element; (344..281);
92% identity.-(5611..5492) W62177 m87d04.r1 Soares mouse
embryo NbME13.5 14.5 Mus musculus cDNA clone 375367.5'
similar to WP:116H12.5 CE00510 KRUPPEL-LIKE ZINC FINGER
PROTEIN ; (35..153); 84% identity.-(5555..5492) T51086
yb71905.r1 Homo sapiens cDNA clone 76664.5'; (1..60); 100%
identity.
```

```
repeat_region complement(5933..6068)
/rpt_family="(TGAA)n"
repeat_region 6098..6204
/rpt_family="(CARG)n"
complement(6191..6312)
repeat_region /rpt_family="(TGAA)n"
7533..7826
/rpt_family="AluJo"
7827..8136
repeat_region /rpt_family="Alusx"
8169..8374
/rpt_family="MIR1C"
8848..9143
/rpt_family="L1MB5"
9150..9364
repeat_region /rpt_family="Alusq"
9366..9655
/rpt_family="Aluv"
complement(9690..9875)
/rpt_family="MER21B"
complement(9892..10186)
/rpt_family="Alusx"
complement(10229..10508)
/rpt_family="AluJo"
complement(10509..10740)
/rpt_family="MER21B"
10741..11044
/rpt_family="Aluv"
11217..11333
/rpt_family="AluJo"
11340..11635
/rpt_family="Alusq"
11636..11937
repeat_region /rpt_family="AlusP"
11954..12436
/rpt_family="MER4A2"
complement(12437..12744)
/rpt_family="AlusP"
12746..12801
repeat_region /rpt_family="MER4A"
12802..12841
/rpt_family="L1PA15"
12845..12973
/rpt_family="AluJo/FLAM"
complement(12974..13267)
/rpt_family="Aluv"
complement(13289..13453)
/rpt_family="L1MC/D"
complement(13471..13647)
/rpt_family="AluJo"
13649..13946
repeat_region /rpt_family="Alusq"
13963..14253
/rpt_family="AluJo"
complement(14292..14463)
/rpt_family="MER3"
14535..14674
repeat_region /rpt_family="Alusq/x"
complement(14779..14839)
/rpt_family="MER58A"
complement(14840..15005)
/rpt_family="Alusq"
complement(15006..15303)
/rpt_family="Alusq"
complement(15306..15438)
/rpt_family="Alusq"
complement(15439..15599)
/rpt_family="MER58B"
complement(15613..15701)
/rpt_family="FLAM.A"
15702..15994
repeat_region /rpt_family="Alusx"
complement(16155..16447)
```

```

ERFQRRIKLGVTQADVGSALANIKIGVGLSGTICRFESLTLSHNNMIALKPIILQ
AMLEAAGAOERKKNKPELFNPGKKRRRSIAAPERKSLAAYAVQPPRESSEKIAAH
AEKIDLKKNVRYVFCNRQKQKMKESAT"
exon
1. .>1149
/gene="Brn-3a"
misc_feature
7. .86
/gene="Brn-3a"
misc_feature
785. .1125
/note="upstream homology domain (UHD)"
/gene="Brn-3a"
/note="POU domain"
BASE COUNT 179 a 416 c 424 g 130 t
ORIGIN
Query Match 9.7%; Score 47.8; DB 9; Length 1149;
Best Local Similarity 50.2%; Pred. No. 98;
Matches 118; Conservative 0; Mismatches 117; Indels 0; Gaps 0;
OY 2 CTCGGGATCATCTTCACCATCGGGCGGAACTCCTGCGGATATCTCTCTCTC 61
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 169 CTGGCGGACACACACACACACACACACACACACACACAGGCGCTCGAACC GGCCATCTG 228
62 CACCGGACCCCCCATGTAGCGGCGCAGCTCGGGCCCTGCGTGGAAAGCTGTACATGCTG 121
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 229 GACCAATCTCTCTCGCCCTGCTCGCGCTCATAGCGCGCGCGCGCGCGCGCGCG 288
122 ATCGGCGGCGTGTGTCGCGCGCGCGCGGCTCTTCCGCTGCTGAGCGATGCCGCTCGTC 181
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 289 GCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 348
182 GCGCTTGGCGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 236
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 349 GCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 403

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2002, 06:59:43 ; Search time 2640 Seconds

(without alignments)
143.309 Million cell updates/sec

Title: US-09-548-449-8

Perfect score: 13

Sequence: 1 aatcaannantta 13

Scoring table:

IDENTITY_MUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 995600

Minimum DB seq length: 0

Maximum DB seq length: 1000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb_ba: *
2: gb_htg: *
3: gb_in: *
4: gb_cm: *
5: gb_ov: *
6: gb_pat: *
7: gb_ph: *
8: gb_pl: *
9: gb_pr: *
10: gb_ro: *
11: gb_sts: *
12: gb_sy: *
13: gb_un: *
14: gb_vl: *
15: em_ba: *
16: em_fun: *
17: em_hum: *
18: em_in: *
19: em_mu: *
20: em_om: *
21: em_or: *
22: em_ov: *
23: em_pat: *
24: em_ph: *
25: em_pl: *
26: em_ro: *
27: em_sts: *
28: em_un: *
29: em_vl: *
30: em_htg_hum: *
31: em_htg_low: *
32: em_htg_other: *
33: em_htg_mus: *
34: em_htg_pin: *
35: em_htg_rod: *
36: em_htg_mam: *
37: em_htg_vrt: *
38: em_sy: *
39: em_htgo_hum: *
40: em_htgo_mus: *
41: em_htgo_other: *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10	76.9	19	6	AX132230 Sequence
2	10	76.9	24	6	AX6535 Sequence 27
3	10	76.9	24	6	AX226158 Sequence
4	10	76.9	28	6	AX060375 Sequence
5	10	76.9	28	6	AX085439 Sequence
6	10	76.9	29	6	AR210200 Sequence
7	10	76.9	29	6	AX402628 Sequence
8	10	76.9	36	6	AR053365 Sequence
9	10	76.9	36	6	AR053365 Sequence
10	10	76.9	36	6	AR131089 Sequence
11	10	76.9	36	6	AR131089 Sequence
12	10	76.9	42	6	A04434 Sequence
13	10	76.9	42	6	A05096 Sequence
14	10	76.9	50	6	AX204152 Sequence
15	10	76.9	50	6	AX204153 Sequence
16	10	76.9	60	3	PFAPPS16DD
17	10	76.9	62	6	AX208903 Sequence
18	10	76.9	65	6	AX482817 Sequence
19	10	76.9	65	6	AX482818 Sequence
20	10	76.9	65	6	AX482861 Sequence
21	10	76.9	65	6	AX483043 Sequence
22	10	76.9	65	6	AX483195 Sequence
23	10	76.9	65	6	AX484722 Sequence
24	10	76.9	65	6	AX484855 Sequence
25	10	76.9	65	6	AX485630 Sequence
26	10	76.9	77	8	QPE489833
27	10	76.9	87	6	AX241028 Sequence
28	10	76.9	100	6	A43843 Sequence
29	10	76.9	100	6	I17420 Sequence 17
30	9	69.2	15	6	I17760 Sequence 46
31	9	69.2	17	6	A97808 Sequence 85
32	9	69.2	17	6	AX263688 Sequence
33	9	69.2	17	6	AX263689 Sequence
34	9	69.2	17	6	AX428686 Sequence
35	9	69.2	18	6	A97809 Sequence
36	9	69.2	18	6	AR195307 Sequence
37	9	69.2	18	6	AX191824 Sequence
38	9	69.2	18	6	AX428687 Sequence
39	9	69.2	18	6	E51089 Sequence
40	9	69.2	19	6	A97810 Sequence
41	9	69.2	19	6	A97929 Sequence 87
42	9	69.2	19	6	AX412034 Sequence 20
43	9	69.2	19	6	AX428688 Sequence
44	9	69.2	19	6	AX428807 Sequence
45	9	69.2	20	6	AR089280 Sequence

ALIGNMENTS

RESULT 1
LOCUS AX132230 19 bp DNA
DEFINITION Sequence 3448 from Patent WO0130362.
ACCESSION AX132230
VERSION AX132230.1 GI:14138535
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 19)
AUTHORS Robbins, J.M. and Trütz, R.
TITLE Ribozyme therapy for the treatment of proliferative skin and eye diseases

Pred. No. is the number of results predicted by chance to have a

JOURNAL Patent: WO 0130362-A 3448 03-MAY-2001;
IMMUSOL, INC. (US)
FEATURES Location/Qualifiers
source 1..19
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Cyclin B1 ribozyme binding site"
BASE COUNT 8 a 2 c 4 g 5 t
ORIGIN

Query Match 76.9%; Score 10; DB 6; Length 19;
Best Local Similarity 76.9%; Pred. No. 1.9e+05;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 AATCAANNATTA 13
||||| |
Db 7 AATCAAGACTTA 19

RESULT 2
A64535 24 bp DNA linear PAT 29-MAR-1999
LOCUS Sequence 27 from Patent W09726331.
DEFINITION A64535
ACCESSION A64535
VERSION A64535.1 GI:3717932
KEYWORDS
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 24)
AUTHORS Kornelik,R.G., Mackenzie,A.E., Roy,N., Robertson,G. and Tamal,K.
TITLE USE OF NEURONAL APOPTOSIS INHIBITOR PROTEIN (NIMP)
JOURNAL Patent: WO 9726331-A 27 24-JUL-1997;
UNIV OTTAWA (CA)
COMMENT Other publication AU 1614997 19970811.
FEATURES Location/Qualifiers
source 1..24
/organism="unidentified"
/db_xref="taxon:32644"
ORIGIN

BASE COUNT 10 a 3 c 2 g 9 t

Query Match 76.9%; Score 10; DB 6; Length 24;
Best Local Similarity 76.9%; Pred. No. 1.9e+05;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 AATCAANNATTA 13
||||| |
Db 8 AATCAATAGTTA 20

RESULT 3
AX226158 24 bp DNA linear PAT 10-SEP-2001
LOCUS Sequence 77 from Patent W00160856.
DEFINITION AX226158
ACCESSION AX226158
VERSION AX226158.1 GI:15555470
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 24)
AUTHORS Vikkula,M.
TITLE vmgjom gene and its mutations causing disorders with a vascular
JOURNAL Patent: WO 0160856-A 77 23-AUG-2001;
UNIVERSITE CATHOLIQUE DE LOUVAIN (BE)
FEATURES Location/Qualifiers
source 1..24
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Oligonucleotide"
ORIGIN

BASE COUNT 11 a 2 c 4 g 7 t

ORIGIN

Query Match 76.9%; Score 10; DB 6; Length 24;
Best Local Similarity 76.9%; Pred. No. 1.9e+05;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 AATCAANNATTA 13
||||| |
Db 9 AATCAAGACTTA 21

RESULT 4
AX060375 28 bp DNA linear PAT 22-JAN-2001
LOCUS Sequence 16 from Patent W00100819.
DEFINITION AX060375
ACCESSION AX060375
VERSION AX060375.1 GI:12405861
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 28)
AUTHORS Chauhan,S.
TITLE Targeted gene replacements in enteric bacteria using linear dna
JOURNAL Patent: WO 0100819-A 16 04-JAN-2001;
E.I. DU PONT DE NEMOURS AND COMPANY (US); GENENCOR INTERNATIONAL,
INC. (US)
FEATURES Location/Qualifiers
source 1..28
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="primer"
ORIGIN

BASE COUNT 6 a 3 c 4 g 15 t

Query Match 76.9%; Score 10; DB 6; Length 28;
Best Local Similarity 76.9%; Pred. No. 1.8e+05;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 AATCAANNATTA 13
||||| |
Db 14 AATCAAGTATTA 2

RESULT 5
AX085439 28 bp DNA linear PAT 09-MAR-2001
LOCUS Sequence 52 from Patent W00112833.
DEFINITION AX085439
ACCESSION AX085439
VERSION AX085439.1 GI:13275494
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 28)
AUTHORS Emplage,M., Haynie,S., Laffend,L., Puccel,J. and Whited,G.
TITLE Process for the biological production of 1,3-propanediol with high
JOURNAL Patent: WO 0112833-A 52 22-FEB-2001;
E.I. DU PONT DE NEMOURS AND COMPANY (US); GENENCOR INTERNATIONAL,
INC. (US)
FEATURES Location/Qualifiers
source 1..28
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="primer-primer"
ORIGIN

BASE COUNT 6 a 3 c 4 g 15 t

Query Match 76.9%; Score 10; DB 6; Length 28;
Best Local Similarity 76.9%; Pred. No. 1.8e+05;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AATCAANNANTTA 13
Db 14 AATCAAGTAATTA 2

RESULT 6
LOCUS AR210200/c 29 bp DNA
DEFINITION Sequence 112 from patent US 6387652..
ACCESSION AR210200
VERSION AR210200.1 GI:21512370
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 29)
AUTHORS Haugland,R. and Vesper,S.
TITLE Method of identifying and quantifying specific fungi and bacteria
JOURNAL Patent: US 6387652-A 112 14-MAY-2002;
FEATURES
source Location/Qualifiers
1..29
/organism="unknown"

BASE COUNT 7 a 0 c 7 g 15 t
ORIGIN

Query Match 76.9%; Score 10; DB 6; Length 29;
Best Local Similarity 76.9%; Pred. No. 1.8e+05;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AATCAANNANTTA 13
Db 27 AATCAATCAATTA 15

RESULT 7
LOCUS AX402628/c 29 bp DNA
DEFINITION Sequence 112 from Patent W00196612.
ACCESSION AX402628
VERSION AX402628.1 GI:21387619
KEYWORDS
SOURCE Paecilomyces variotii.
ORGANISM Paecilomyces variotii.
REFERENCE 1
AUTHORS Haugland,R. and Vesper,S.
TITLE Method of identifying and quantifying specific fungi and bacteria
JOURNAL Patent: WO 0196612-A 112 20-DEC-2001;
FEATURES
source Location/Qualifiers
1..29
/organism="Paecilomyces variotii"
/db_xref="taxon:45396"

BASE COUNT 7 a 0 c 7 g 15 t
ORIGIN

Query Match 76.9%; Score 10; DB 6; Length 29;
Best Local Similarity 76.9%; Pred. No. 1.8e+05;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AATCAANNANTTA 13
Db 27 AATCAATCAATTA 15

RESULT 8
LOCUS AR053365 36 bp DNA
DEFINITION Sequence 1 from patent US 5834236.
ACCESSION AR053365
VERSION AR053365.1 GI:5978227

KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 36)
AUTHORS Lamb,C.J., Doerner,P. and Laible,G.
TITLE AATT repeat transcription enhancer element
JOURNAL Patent: US 5834236-A 1 10-NOV-1998;
FEATURES
source Location/Qualifiers
1..36
/organism="unknown"

BASE COUNT 16 a 1 c 2 g 17 t
ORIGIN

Query Match 76.9%; Score 10; DB 6; Length 36;
Best Local Similarity 76.9%; Pred. No. 1.7e+05;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AATCAANNANTTA 13
Db 9 AATCAATTAATTA 21

RESULT 9
LOCUS AR053365/c 36 bp DNA
DEFINITION Sequence 1 from patent US 5834236.
ACCESSION AR053365
VERSION AR053365.1 GI:5978227
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 36)
AUTHORS Lamb,C.J., Doerner,P. and Laible,G.
TITLE AATT repeat transcription enhancer element
JOURNAL Patent: US 5834236-A 1 10-NOV-1998;
FEATURES
source Location/Qualifiers
1..36
/organism="unknown"

BASE COUNT 16 a 1 c 2 g 17 t
ORIGIN

Query Match 76.9%; Score 10; DB 6; Length 36;
Best Local Similarity 76.9%; Pred. No. 1.7e+05;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AATCAANNANTTA 13
Db 32 AATCAATTAATTA 20

RESULT 10
LOCUS AR131089 36 bp DNA
DEFINITION Sequence 1 from patent US 6191258.
ACCESSION AR131089
VERSION AR131089.1 GI:14119414
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 36)
AUTHORS Lamb,C.J., Doerner,P. and Laible,G.
TITLE Purified palindromic element binding factor
JOURNAL Patent: US 6191258-A 1 20-FEB-2001;
FEATURES
source Location/Qualifiers
1..36
/organism="unknown"

BASE COUNT 16 a 1 c 2 g 17 t
ORIGIN

Query Match 76.9%; Score 10; DB 6; Length 36;

Best Local Similarity 76.9%; Pred. No. 1.7e+05;
Matches 10; Conservative 0; Mismatches 3; Indels 0;
QY 1 AATCAANNNTTA 13
||||| 1111
DB 9 AATCAATTAATTA 21

RESULT 11
ARI31089/c ARI31089 36 bp DNA linear PAT 16-MAY-2001
LOCUS DEFINITION Sequence 1 from patent US 6191258.
ACCESSION ARI31089
VERSION ARI31089.1 GI:14119414
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 36)
AUTHORS Lamb,C.J., Doerner,P. and Lalble,G.
TITLE Purified palindromic element binding factor
JOURNAL Patent: US 6191258-A 1 20-FEB-2001;
FEATURES Location/Qualifiers
source 1..36

BASE COUNT 16 a 1 c 2 g 17 t
ORIGIN

Query Match 76.9%; Score 10; DB 6; Length 36;
Best Local Similarity 76.9%; Pred. No. 1.7e+05;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AATCAANNNTTA 13
||||| 1111
DB 32 AATCAATTAATTA 20

RESULT 12
A04434/c A04434 42 bp DNA linear PAT 29-APR-1993
LOCUS DEFINITION Oligonucleotide U9 for salmon growth hormone.
ACCESSION A04434
VERSION A04434.1 GI:344937
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
FEATURES Location/Qualifiers
source 1..42
/organism="synthetic construct"
/db_xref="taxon:32630"

BASE COUNT 8 a 6 c 10 g 18 t
ORIGIN

Query Match 76.9%; Score 10; DB 6; Length 42;
Best Local Similarity 76.9%; Pred. No. 1.7e+05;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AATCAANNNTTA 13
||||| 1111
DB 20 AATCAACAAGTTA 8

RESULT 13
A05096 A05096 42 bp DNA linear PAT 30-APR-1993
LOCUS DEFINITION Oligonucleotide L9 for salmon growth hormone.
ACCESSION A05096
VERSION A05096.1 GI:344983
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
FEATURES Location/Qualifiers
source 1..50

FEATURES Location/Qualifiers
source 1..42
/organism="synthetic construct"
/db_xref="taxon:32630"
BASE COUNT 17 a 9 c 6 g 10 t
ORIGIN

Query Match 76.9%; Score 10; DB 6; Length 42;
Best Local Similarity 76.9%; Pred. No. 1.7e+05;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AATCAANNNTTA 13
||||| 1111
DB 4 AATCAACAAGTTA 16

RESULT 14
AX204152 AX204152 50 bp DNA linear PAT 30-AUG-2001
LOCUS DEFINITION Sequence 258 from Patent WO0148245.
ACCESSION AX204152
VERSION AX204152.1 GI:15393650
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 50)
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Shimketa,R.A. and Leach,M.
Nucleic acids containing single nucleotide polymorphisms and
METHODS methods of use thereof
JOURNAL Patent: WO 0148245-A 258 05-JUL-2001;
FEATURES Location/Qualifiers
source 1..50
/organism="Homo sapiens"
/db_xref="taxon:9606"

misc_feature 25..26
/note="Nucleotide deleted between bases 25 and 26
Accession number cg43967861"
variation 26
/note="single nucleotide polymorphism"
BASE COUNT 20 a 14 c 3 g 13 t
ORIGIN

Query Match 76.9%; Score 10; DB 6; Length 50;
Best Local Similarity 76.9%; Pred. No. 1.6e+05;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AATCAANNNTTA 13
||||| 1111
DB 24 AATCAACAAGTTA 36

RESULT 15
AX204153 AX204153 50 bp DNA linear PAT 30-AUG-2001
LOCUS DEFINITION Sequence 259 from Patent WO0148245.
ACCESSION AX204153
VERSION AX204153.1 GI:15393651
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 50)
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Shimketa,R.A. and Leach,M.
Nucleic acids containing single nucleotide polymorphisms and
METHODS methods of use thereof
JOURNAL Patent: WO 0148245-A 259 05-JUL-2001;
FEATURES Location/Qualifiers
source 1..50

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
misc_feature 25..26
    /note="Nucleotide deleted between bases 25 and 26"
    Accession number cg43967861"
variation 26
    /note="single nucleotide polymorphism"
BASE COUNT 20 a 13 c 4 g 13 t
ORIGIN
Query Match 76.9%; Score 10; DB 6; Length 50;
Best local Similarity 76.9%; Pred. No. 1.6e+05;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AATCAANNANTTA 13
    ||||| 1 |||
Db 23 AATCAACAATTA 35

```

Search completed: December 26, 2002, 13:25:58
 Job time : 2652 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2002, 08:20:51 ; Search time 54 Seconds
(without alignments)
97.797 Million cell updates/sec

Title: US-09-548-449-8
Perfect score: 13
Sequence: 1 aatcaannanta 13

Scoring table: IDENTITY_NDC
Gapop 10.0 , Gapext 1.0

Searched: 363474 seqs, 203117208 residues

Total number of hits satisfying chosen parameters: 221248

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA.*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10	76.9	24	8	US-08-913-322-27
2	10	76.9	36	10	US-09-839-743-1
3	10	76.9	36	10	US-09-839-743-1
4	10	76.9	44	10	US-09-839-743-28
5	10	76.9	44	10	US-09-839-743-28
6	10	76.9	62	10	US-09-777-564-743
7	10	76.9	65	10	US-09-783-590-1151
8	10	76.9	59	10	US-09-969-373-1158
9	9.6	73.8	78	10	US-09-764-877-445
9	9.6	73.8	20	8	US-08-983-605-453
10	9	69.2	20	10	US-09-758-881-20
11	9	69.2	20	10	US-09-452-599-39
12	9	69.2	22	9	US-09-835-371-46
13	9	69.2	22	10	US-09-860-784-25
14	9	69.2	24	9	US-10-014-436-10
15	9	69.2	24	10	US-09-969-373-3966
16	9	69.2	26	10	US-09-837-235-35
17	9	69.2	26	10	US-09-837-235-36
18	9	69.2	26	10	US-09-837-235-36
19	9	69.2	26	10	US-09-775-879-2

20	9	69.2	26	10	US-09-860-784-97	Sequence 97, Appl
21	9	69.2	26	10	US-09-860-784-98	Sequence 98, Appl
22	9	69.2	28	9	US-10-037-598-15	Sequence 15, Appl
23	9	69.2	33	10	US-09-955-462A-3	Sequence 3, Appl
24	9	69.2	40	9	US-09-832-292-25	Sequence 25, Appl
25	9	69.2	53	10	US-09-732-914-128	Sequence 128, Appl
26	9	69.2	64	10	US-09-865-807-50	Sequence 50, Appl
27	9	69.2	65	10	US-09-878-574-10483	Sequence 10483, A
28	9	69.2	67	8	US-08-808-031A-26	Sequence 26, Appl
29	9	69.2	68	10	US-09-983-965-154	Sequence 154, Appl
30	9	69.2	72	9	US-09-771-382-48	Sequence 48, Appl
31	9	69.2	79	10	US-09-923-876-4907	Sequence 4907, Ap
32	9	69.2	83	10	US-09-864-761-18906	Sequence 18906, A
33	9	69.2	83	10	US-09-815-242-787	Sequence 787, Appl
34	9	69.2	93	10	US-09-971-798-3	Sequence 3, Appl
35	9	69.2	96	10	US-09-867-701-1092	Sequence 1092, Ap
36	9	69.2	97	10	US-09-294-093B-5580	Sequence 5580, Ap
37	9	69.2	97	10	US-09-869-373-467	Sequence 467, Appl
38	9	69.2	99	10	US-09-871-798-20	Sequence 20, Appl
39	9	69.2	99	10	US-09-971-798-20	Sequence 21, Appl
40	9	69.2	100	10	US-09-728-445-116	Sequence 116, Appl
41	8.4	64.6	15	10	US-09-504-231A-611	Sequence 611, Appl
42	8.4	64.6	15	10	US-09-274-553D-611	Sequence 611, Appl
43	8.4	64.6	20	9	US-09-774-223-4	Sequence 4, Appl
44	8.4	64.6	21	10	US-09-920-342-6	Sequence 6, Appl
45	8.4	64.6	22	8	US-08-913-322-26	Sequence 26, Appl

ALIGNMENTS

RESULT 1
US-08-913-322-27
Sequence 27, Application US/08913322
Patent No. US20020137028A1
GENERAL INFORMATION:
APPLICANT: Korneluk, Robert G.
APPLICANT: Mackenzie, Alexander E.
APPLICANT: Roy, Natalie
APPLICANT: Robertson, George
APPLICANT: Tamaki, Katsu
TITLE OF INVENTION: USER OF NEURONAL APOPTOSIS INHIBITOR
TITLE OF INVENTION: (NAIP)
FILE REFERENCE: 07891/013001
CURRENT APPLICATION NUMBER: US/08/913, 322
CURRENT FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: PCT/IB97/00142
EARLIER FILING DATE: 1997-01-17
EARLIER APPLICATION NUMBER: GB 9601108.5
NUMBER OF SEQ ID NOS: 27
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 27
LENGTH: 24
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic primer based on Homo sapiens
US-08-913-322-27

Query Match 76.9% Score 10; DB 8; Length 24;
Best Local Similarity 76.9% Pred. No. 1.9e+03;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AATCAANNANTTA 13
DB 8 AATCAATTAAGTTA 20

RESULT 2
US-09-839-743-1
Sequence 1, Application US/09839743
Patent No. US20020146824A1

```
GENERAL INFORMATION:
APPLICANT: The Salk Institute for Biological Sciences
APPLICANT: Lamb, Christopher
APPLICANT: Doerner, Peter
APPLICANT: Laidle, Goetz
TITLE OF INVENTION: Transcription Factor and Methods of Use Thereof
FILE REFERENCE: SALKINS.008DV3
CURRENT APPLICATION NUMBER: US/09/839,743
PRIOR FILING DATE: 2001-04-19
PRIOR APPLICATION NUMBER: US 09/401,336
PRIOR FILING DATE: 1999-09-21
PRIOR APPLICATION NUMBER: US 09/189,344
PRIOR FILING DATE: 1998-11-10
PRIOR APPLICATION NUMBER: US 08/669,721
PRIOR FILING DATE: 1996-06-27
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 36
TYPE: DNA
ORGANISM: Nicotiana tabacum
US-09-839-743-1
```

```
Query Match      76.9%; Score 10; DB 10; Length 36;
Best Local Similarity 76.9%; Pred. No. 2e+03;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
OY 1 AATCAANNANTTA 13
    ||||| 1 |||
DB 9 AATCAATTATTA 21
```

```
RESULT 3
US-09-839-743-1/c
Sequence 1, Application US/09839743
Patent No. US20020146824A1
GENERAL INFORMATION:
APPLICANT: The Salk Institute for Biological Sciences
APPLICANT: Lamb, Christopher
APPLICANT: Doerner, Peter
APPLICANT: Laidle, Goetz
TITLE OF INVENTION: No. US20020146824A1 Transcription Enhancer Element and
TITLE OF INVENTION: Transcription Factor and Methods of Use Thereof
FILE REFERENCE: SALKINS.008DV3
CURRENT APPLICATION NUMBER: US/09/839,743
CURRENT FILING DATE: 2001-04-19
PRIOR APPLICATION NUMBER: US 09/401,336
PRIOR FILING DATE: 1999-09-21
PRIOR APPLICATION NUMBER: US 09/189,344
PRIOR FILING DATE: 1998-11-10
PRIOR APPLICATION NUMBER: US 08/669,721
PRIOR FILING DATE: 1996-06-27
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 36
TYPE: DNA
ORGANISM: Nicotiana tabacum
US-09-839-743-1
```

```
Query Match      76.9%; Score 10; DB 10; Length 36;
Best Local Similarity 76.9%; Pred. No. 2e+03;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
OY 1 AATCAANNANTTA 13
    ||||| 1 |||
DB 32 AATCAATTATTA 20
```

```
RESULT 4
US-09-839-743-28
Sequence 28, Application US/09839743
```

```
Patent No. US20020146824A1
GENERAL INFORMATION:
APPLICANT: The Salk Institute for Biological Sciences
APPLICANT: Lamb, Christopher
APPLICANT: Doerner, Peter
APPLICANT: Laidle, Goetz
TITLE OF INVENTION: Transcription Factor and Methods of Use Thereof
FILE REFERENCE: SALKINS.008DV3
CURRENT APPLICATION NUMBER: US/09/839,743
CURRENT FILING DATE: 2001-04-19
PRIOR APPLICATION NUMBER: US 09/401,336
PRIOR FILING DATE: 1999-09-21
PRIOR APPLICATION NUMBER: US 09/189,344
PRIOR FILING DATE: 1998-11-10
PRIOR APPLICATION NUMBER: US 08/669,721
PRIOR FILING DATE: 1996-06-27
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 28
LENGTH: 44
TYPE: DNA
ORGANISM: Nicotiana tabacum
US-09-839-743-28
```

```
Query Match      76.9%; Score 10; DB 10; Length 44;
Best Local Similarity 76.9%; Pred. No. 2e+03;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
OY 1 AATCAANNANTTA 13
    ||||| 1 |||
DB 2 AATCAATTATTA 14
```

```
RESULT 5
US-09-839-743-28/c
Sequence 28, Application US/09839743
Patent No. US20020146824A1
GENERAL INFORMATION:
APPLICANT: The Salk Institute for Biological Sciences
APPLICANT: Lamb, Christopher
APPLICANT: Doerner, Peter
APPLICANT: Laidle, Goetz
TITLE OF INVENTION: No. US20020146824A1 Transcription Enhancer Element and
TITLE OF INVENTION: Transcription Factor and Methods of Use Thereof
FILE REFERENCE: SALKINS.008DV3
CURRENT APPLICATION NUMBER: US/09/839,743
CURRENT FILING DATE: 2001-04-19
PRIOR APPLICATION NUMBER: US 09/401,336
PRIOR FILING DATE: 1999-09-21
PRIOR APPLICATION NUMBER: US 09/189,344
PRIOR FILING DATE: 1998-11-10
PRIOR APPLICATION NUMBER: US 08/669,721
PRIOR FILING DATE: 1996-06-27
NUMBER OF SEQ ID NOS: 28
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 28
LENGTH: 44
TYPE: DNA
ORGANISM: Nicotiana tabacum
US-09-839-743-28
```

```
Query Match      76.9%; Score 10; DB 10; Length 44;
Best Local Similarity 76.9%; Pred. No. 2e+03;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
OY 1 AATCAANNANTTA 13
    ||||| 1 |||
DB 29 AATCAATTATTA 17
```

```
RESULT 6
US-09-777-564-743/c
```



```
; Sequence 743, Application US/09777564
; Patent No. US20020022591A1
; GENERAL INFORMATION:
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannon, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.493
; CURRENT APPLICATION NUMBER: US/09/777,564
; CURRENT FILING DATE: 2001-02-05
; NUMBER OF SEQ ID NOS: 1730
; SOFTWARE: FastSeq for Window Version 4.0
; SEQ ID NO 743
; LENGTH: 62
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-777-564-743

Query Match          76.9%; Score 10; DB 10; Length 62;
Best Local Similarity 76.9%; Pred. No. 2.1e+03;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 AATCAANNANTTA 13
        ||||| | |||
Db      57 AATCAATAATTTA 45

RESULT 7
US-09-783-590-1151
; Sequence 1151, Application US/09783590
; Patent No. US20020110850A1
; GENERAL INFORMATION:
; APPLICANT: Dillon, Patrick J.
; APPLICANT: Haseltine, William A.
; APPLICANT: Li, Haodong
; APPLICANT: Rosen, Craig A.
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Genes, Sequences, and Expression Products 16.2
; FILE REFERENCE: PO-16.201
; CURRENT APPLICATION NUMBER: US/09/783,590
; CURRENT FILING DATE: 2000-02-15
; PRIOR APPLICATION NUMBER: 08/420,856
; PRIOR FILING DATE: 1995-04-12
; PRIOR APPLICATION NUMBER: 08/346,731
; PRIOR FILING DATE: 1994-11-21
; NUMBER OF SEQ ID NOS: 12485
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1151
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-783-590-1151

Query Match          76.9%; Score 10; DB 10; Length 65;
Best Local Similarity 76.9%; Pred. No. 2.1e+03;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 AATCAANNANTTA 13
        ||||| | |||
Db      11 AATCAATAATTTA 23

RESULT 8
US-09-969-373-1158/c
; Sequence 1158, Application US/09969373
; Patent No. US20020133852A1
; GENERAL INFORMATION:
; APPLICANT: Efferetz, Roger J.
; APPLICANT: Hauge, Brian M.

; TITLE OF INVENTION: Soybean SSRs and Methods of Genotyping
; FILE REFERENCE: 38-10(52679)A
; CURRENT APPLICATION NUMBER: US/09/969,373
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US 09/754,853
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 09/760,427
; PRIOR FILING DATE: 2001-01-13
; PRIOR APPLICATION NUMBER: US 09/855,768
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 4593
; SEQ ID NO 1158
; LENGTH: 99
; TYPE: DNA
; ORGANISM: Glycine max
US-09-969-373-1158

Query Match          76.9%; Score 10; DB 10; Length 99;
Best Local Similarity 76.9%; Pred. No. 2.1e+03;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 AATCAANNANTTA 13
        ||||| | |||
Db      45 AATCAAGTAATTA 33

RESULT 9
US-09-764-877-445/c
; Sequence 445, Application US/09764877
; Patent No. US20020147140A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC005
; CURRENT APPLICATION NUMBER: US/09/764,877
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4031
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 445
; LENGTH: 78
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-877-445

Query Match          73.8%; Score 9.6; DB 10; Length 78;
Best Local Similarity 69.2%; Pred. No. 3.6e+03;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      1 AATCAANNANTTA 13
        ||||| | |||
Db      47 AATCAAAAAAATTA 35

RESULT 10
US-08-983-605-453/c
; Sequence 453, Application US/08983605A
; Patent No. US2002006118A1
; GENERAL INFORMATION:
; APPLICANT: Roder, Marion
; TITLE OF INVENTION: Microsatellite Markers for Plants of the Species
; TITLE OF INVENTION: Trifolium aestivum and Tribe Trifolaceae and the Use of
; FILE REFERENCE: 2936.10400
; CURRENT APPLICATION NUMBER: US/08/983,605A
; CURRENT FILING DATE: 1998-05-01
; EARLIER APPLICATION NUMBER: DE 195 25 284.5
; PRIOR FILING DATE: 1995-06-28
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 453
; LENGTH: 20
; TYPE: DNA
```

ORGANISM: Trifolium aestivum
US-08-983-605-453

Query Match 69.2%; Score 9; DB 8; Length 20;
Best Local Similarity 75.0%; Pred. No. 7.5e+03;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 AATCAANNANTT 12
|||||
DB 15 AATCAATAAGTT 4

RESULT 11

US-09-758-881-20/c
Sequence 20, Application US/09758881
Patent No. US20010029250A1
GENERAL INFORMATION:
APPLICANT: Karitas, James G
TITLE OF INVENTION: Antisense Oligonucleotide Modulation of STAT3
FILE REFERENCE: ISPH-0532
CURRENT APPLICATION NUMBER: US/09/758,881
CURRENT FILING DATE: 2001-01-11
PRIOR APPLICATION NUMBER: PCT/US00/09054
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: 09/288,461
PRIOR FILING DATE: 1999-04-08
NUMBER OF SEQ ID NOS: 152
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 20
LENGTH: 20
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-758-881-20

Query Match 69.2%; Score 9; DB 10; Length 20;
Best Local Similarity 75.0%; Pred. No. 7.5e+03;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 AATCAANNANTT 12
|||||
DB 20 AATCAAGCAGTT 9

RESULT 12

US-09-452-599-39
Sequence 39, Application US/09452599
Patent No. US20020055101A1
GENERAL INFORMATION:
APPLICANT: Bergeron, Michel G.
APPLICANT: Ouellette, Marc
APPLICANT: Roy, Paul H.
TITLE OF INVENTION: Specific and Universal Probes and Amplification Primers
TITLE OF INVENTION: to Rapidly Detect and Identify Common Bacterial
TITLE OF INVENTION: Pathogens and Antibiotic Resistance Genes from Clinical
FILE REFERENCE: 12287,31
CURRENT APPLICATION NUMBER: US/09/452,599
CURRENT FILING DATE: 1999-12-01
PRIOR APPLICATION NUMBER: 08/526,840
PRIOR FILING DATE: 1995-09-11
PRIOR APPLICATION NUMBER: 08/304,732
PRIOR FILING DATE: 1994-09-12
NUMBER OF SEQ ID NOS: 177
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 39
LENGTH: 20
TYPE: DNA
ORGANISM: Enterococcus faecalis
US-09-452-599-39

Query Match 69.2%; Score 9; DB 10; Length 20;
Best Local Similarity 75.0%; Pred. No. 7.5e+03;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 AATCAANNNTTA 13
|||||
DB 5 AATCAACATTA 16

RESULT 13

US-09-835-371-46
Sequence 46, Application US/09835371
Publication No. US20020187473A1
GENERAL INFORMATION:
APPLICANT: UHLMANN, Eugen
APPLICANT: BREIPOHL, Gerhard
APPLICANT: WILL, David W
TITLE OF INVENTION: POLYAMIDE NUCLEIC ACID DERIVATIVES, AND AGENTS AND
TITLE OF INVENTION: PROCESSES FOR PREPARING THEM
FILE REFERENCE: 02481,1743 SEQUENCE LISTING
CURRENT APPLICATION NUMBER: US/09/835,371
CURRENT FILING DATE: 2001-04-17
NUMBER OF SEQ ID NOS: 53
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 46
LENGTH: 22
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: base sequence
US-09-835-371-46

Query Match 69.2%; Score 9; DB 9; Length 22;
Best Local Similarity 75.0%; Pred. No. 7.5e+03;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 AATCAANNANTT 12
|||||
DB 2 AATCAATGACTT 13

RESULT 14

US-09-860-784-25
Sequence 25, Application US/09860784
Patent No. US20020151512A1
GENERAL INFORMATION:
APPLICANT: PEYMAN, Anuschirwan
UHLMANN, Eugen
TITLE OF INVENTION: G CAP-STABILIZED OLIGONUCLEOTIDES
NUMBER OF SEQUENCES: 105
CORRESPONDENCE ADDRESS:
ADDRESS: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/860,784
FILING DATE: 21-May-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/594,452
FILING DATE: 04-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: SANDERCOCK, COLIN G.
REGISTRATION NUMBER: 31,298

REFERENCE/DOCKET NUMBER: 18748/264/HOCE
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-860-784-25
SEQUENCE DESCRIPTION: SEQ ID NO: 25:
US-09-860-784-25

Query Match 69.2%; Score 9; DB 10; Length 22;
Best Local Similarity 75.0%; Pred. No. 7.5e+03;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AATCAANNANTT 12
||||| 1 1
Db 2 AATCAATGACTT 13

RESULT 15
US-10-014-436-10
Sequence 10, Application US/10014436
Publication No. US20020182699A1
GENERAL INFORMATION:
APPLICANT: HATADA, Yuji
APPLICANT: IGARASHI, Kazuaki
APPLICANT: OZAKI, Katsuya
APPLICANT: AWA, Katsutoshi
APPLICANT: KAWAI, Shuji
APPLICANT: ITO, Susumu
TITLE OF INVENTION: Gene for Enzyme Having Both Alkaline Pullulanase and Alkaline Alp
FILE REFERENCE: 2173-0122P
CURRENT APPLICATION NUMBER: US/10/014,436
CURRENT FILING DATE: 2002-05-21
PRIOR APPLICATION NUMBER: JP 111547/1995
PRIOR FILING DATE: 1995-05-10
PRIOR APPLICATION NUMBER: PCT/JP96/01243
PRIOR FILING DATE: 1996-05-10
PRIOR APPLICATION NUMBER: US 08/952,084
PRIOR FILING DATE: 1997-11-10
PRIOR APPLICATION NUMBER: US 09/514,302
PRIOR FILING DATE: 2000-02-28
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn version 3.1
SEQ ID NO 10
LENGTH: 24
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Primer 4. Directed to Bacillus sp. used between XbaI to 1.2 kb u
US-10-014-436-10

Query Match 69.2%; Score 9; DB 9; Length 24;
Best Local Similarity 75.0%; Pred. No. 7.6e+03;
Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AATCAANNANTT 12
||||| 1 1
Db 12 AATCAAGAAAT 23

Search completed: December 26, 2002, 14:04:00
Job time : 57 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 25, 2002, 18:16:12 ; Search time 278 Seconds
(without alignments)
3985.551 Million cell updates/sec

Title: US-09-548-449-3
Perfect score: 492
Sequence: 1 actcgcgcgcacatcacc.....gccttaacctagcaagac 492

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 2185239 seqs, 1125999159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	492	100.0	492	21 AAA95502 P. aeruginosa ANR
2	492	100.0	492	21 AAZ58691 Pseudomonas ARN pr
3	492	100.0	492	22 ABA02573 Pseudomonas aerugi
4	47.8	9.7	1272	20 AAX09010 Brn-3a polynucleot
5	47.8	9.7	1272	21 AAA29006 Human transcriptio
6	46.8	9.5	125401	22 AAD17186 Streptomyces ncurs
7	46	9.3	2765	22 AAAH1550 Human Rtl1 gamma n
8	46	9.3	3131	22 AAAH1549 Human Rtl1 beta nu
9	46	9.3	3332	22 AAAH1548 Human Rtl1 alpha'

C	10	46	9.3	3344	22	AAH41505	Human Rtl1 encodin
C	11	45.6	9.3	530	21	AAA95215	Human OCT-1 part1
C	12	45.6	9.3	4524	20	AAV33912	Nucleotide sequenc
C	13	44.2	9.0	1146	16	AAO83793	S. avermitilis BCK
C	14	44.2	9.0	2728	16	AAO83797	S. avermitilis BCK
C	15	44	8.9	1388	23	ABV23346	human prostate exp
C	16	44	8.9	1388	23	ABV29200	human prostate exp
C	17	44	8.9	4257	19	AAV68520	The nucleotide seq
C	18	44	8.9	4257	19	AAV68520	Infected cell prot
C	19	44	8.9	12001	16	AAO76132	HSV L/ST region.
C	20	43.8	8.9	68750	21	AAZ55887	Sorangium cellulos
C	21	43.8	8.9	71989	21	AAA29349	Sorangium cellulos
C	22	43.6	8.9	1424	23	ABL20183	Drosophila melanog
C	23	43.6	8.9	1469	23	ABL20181	Drosophila melanog
C	24	43.4	8.8	1502	23	ABL12317	Drosophila melanog
C	25	43.4	8.8	1131	22	AAH52039	Mycobacterium tube
C	26	43.4	8.8	1305	23	AAH54093	Pseudomonas aerugi
C	27	43.2	8.8	10468	23	ABL12316	Drosophila melanog
C	28	43.2	8.8	10468	23	ABL20180	Drosophila melanog
C	29	43.2	8.8	10468	23	ABL20182	Drosophila melanog
C	30	43.2	8.8	65140	22	AAZ17184	Streptomyces ncurs
C	31	42.8	8.7	367	20	AAZ19390	M. tuberculosis an
C	32	42.8	8.7	367	20	AAZ19178	M. tuberculosis re
C	33	42.8	8.7	1028	13	AAQ27091	XY26 probe. Homo
C	34	42.8	8.7	1028	22	AAH15342	1.0kb PstI fragmen
C	35	42.8	8.7	1028	22	AAH15347	Human Fragile X sy
C	36	42.6	8.7	603	23	AAH78361	DNA encoding novel
C	37	42.4	8.6	2846	19	AAV41260	Chlamydomonas rein
C	38	42.4	8.6	44377	18	AAH78508	Platanolide syntha
C	39	42.4	8.6	44377	18	AAH78508	Platanolide syntha
C	40	42	8.5	509	24	ABK34507	Human cDNA for nov
C	41	42	8.5	567	21	AAA29550	HIV codon altered
C	42	42	8.5	846	22	AAH06566	Human cDNA clone (
C	43	42	8.5	1651	24	AAH77281	Human hepatocyte g
C	44	42	8.5	2601	24	AAH28495	Human extracellular
C	45	42	8.5	3260	22	AAH17836	Human cDNA sequenc

ALIGNMENTS

RESULT 1	AAA95502	standard; DNA; 492 BP.
ID	AAA95502	
AC	AAA95502	
XX	27-FEB-2001	(first entry)
DT		
DE	P. aeruginosa ANR promoter sequence.	
XX		
XX	Bacteria: Infectio; drug-resistant pathogen; cancer; typhoid fever;	
KW	bacterial meningitis; tuberculosis; antisense strand; ribozyme;	
KW	toxic protein; ds.	
XX		
OS	Pseudomonas aeruginosa.	
XX		
PN	WO200061804-A1.	
XX		
PD	19-OCT-2000.	
XX		
PF	14-APR-2000; 2000WO-US10229.	
XX		
PR	14-APR-1999; 99US-0291902.	
XX		
PR	13-APR-2000; 2000US-0548445.	
XX		
PA	(MUSC-) MUSC FOUND RES DEV.	
XX	(UYPE-) UNIV PENN STATE.	
PI	Norris J, Clawson G, Westwater C, Schofield D, Schmidt M, Hoel B;	
XX	Dolan J, Pan W;	
DR	WPI: 2000-638570/61.	

Db 241 TGAAGTCACATTCACAGTCCGCGGAAAAATGGAATTCCTTCATTGATGCGCCAC 300
 QY 301 GCGTCGGCAACTGAGCCCTTTCTGCGCCCTTGACAGGTGCGACAGTAGTCGCA 360
 Db 301 GCGTCGGCAACTGAGCCCTTTCTGCGCCCTTGACAGGTGCGACAGTAGTCGCA 360
 QY 361 GTTGTGTAGCAAGTACATGATTTGGAAGCCATCGGCGCTGTACAGAAATGTTGCTGCC 420
 Db 361 GTTGTGTAGCAAGTACATGATTTGGAAGCCATCGGCGCTGTACAGAAATGTTGCTGCC 420
 QY 421 AGACCTATGCTGCGACCCCGATCGGCGCTGTACCTTACTCTGTTGCTGCTTAA 480
 Db 421 AGACCTATGCTGCGACCCCGATCGGCGCTGTACCTTACTCTGTTGCTGCTTAA 480
 QY 481 CCTAGCAAGGAC 492
 Db 481 CCTAGCAAGGAC 492

RESULT 3
 ABA02573 standard; DNA; 492 BP.
 ABA02573:
 XX
 AC ABA02573:
 XX
 DT 05-FEB-2002 (first entry)
 XX
 DE Pseudomonas aeruginosa promoter ANR.
 XX
 KM Infection; antisense RNA; ribozyme; DNAzyme; antiviral; gene therapy;
 KM papilloma virus; hepatitis B virus; cytotoxic; cytostatic; warty;
 KM cervical dysplasia; cervical carcinoma; carcinoma; laryngeal papilloma;
 KM ANR promoter; ds.
 XX
 OS Pseudomonas aeruginosa.
 XX
 PN WO200179524-A2.
 PD 25-OCT-2001.
 XX
 PF 13-APR-2001; 2001WO-US12130.
 XX
 PR 13-APR-2000; 2000US-0548449.
 PR 07-DEC-2000; 2000US-251810P.
 XX
 PA (UYSC-) UNIV SOUTH CAROLINA.
 PA (PENN-) PENN STATE RES FOUND.
 XX
 PI Norris JS, Clawson GA, Westwater C, Schofield D, Schmidt WG;
 PI Hoel B, Dolan J, Pan W;
 DR WPI; 2001-607700/69.
 XX
 PT Novel nucleic acid for the treatment of papilloma or hepatitis virus
 PT induced conditions comprises a catalytic region which produces a
 PT cytotoxic or cytostatic effect in the infected cell -
 XX
 PS Examples; Fig 1; 143bp; English.
 XX
 CC The invention relates to the discovery, identification and
 CC characterization of toxic agents lethal to pathogens and methods for
 CC targeting such toxic agents to a pathogen or pathogen infected cells in
 CC order to treat and/or eradicate the infection. In particular the
 CC invention relates to at least one nucleic acid molecule, which
 CC specifically hybridizes to mRNA encoding at least one viral protein
 CC associated with the transformation or plasmid copy number control, which
 CC hybridizes to a viral polyadenylation signal or a core, pre core or
 CC polymerase encoding sequence. Specifically, the invention relates to the
 CC delivery of one or more toxic gene products, antisense RNAs, ribozymes,
 CC DNAzymes or a combination thereof. The nucleic acids have antiviral
 CC activity and can be used in gene therapy. They are useful for the
 CC treatment of papilloma or hepatitis virus induced conditions and can
 CC produce a cytotoxic or cytostatic effect in papillomavirus or hepatitis B

CC infected cells. The papilloma virus induced condition is selected from
 CC warts, cervical dysplasia, cervical carcinoma, carcinoma in situ and
 CC laryngeal papilloma. The present sequence is that of a Pseudomonas
 CC aeruginosa promoter sequence, useful to the invention.
 XX
 SQ Sequence 492 BP: 77 A; 164 C; 147 G; 104 T; 0 other:
 Query Match 100.0%; Score 492; DB 22; Length 492;
 Best Local Similarity 100.0%; Pred. No. 3.3e-103;
 Matches 492; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACTGCGGATCATCTTACCATTCGCGGCAATCTCTCGGATATCTGCTCTCTCT 60
 Db 1 ACTGCGGATCATCTTACCATTCGCGGCAATCTCTCGGATATCTGCTCTCTCT 60
 QY 61 CCACCGGACCCCGATGATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
 Db 61 CCACCGGACCCCGATGATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120
 QY 121 GATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
 Db 121 GATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180
 QY 181 CGGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
 Db 181 CGGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
 QY 241 TGAGGTCACATTCACATTCACATTCACATTCACATTCACATTCACATTCACATTC 300
 Db 241 TGAGGTCACATTCACATTCACATTCACATTCACATTCACATTCACATTCACATTC 300
 QY 301 GCGTCGGCAACTGAGCCCTTTCTGCGCCCTTGACAGGTGCGACAGTAGTCGCA 360
 Db 301 GCGTCGGCAACTGAGCCCTTTCTGCGCCCTTGACAGGTGCGACAGTAGTCGCA 360
 QY 361 GTTGTGTAGCAAGTACATGATTTGGAAGCCATCGGCGCTGTACAGAAATGTTGCTGCC 420
 Db 361 GTTGTGTAGCAAGTACATGATTTGGAAGCCATCGGCGCTGTACAGAAATGTTGCTGCC 420
 QY 421 AGACCTATGCTGCGACCCCGATCGGCGCTGTACCTTACTCTGTTGCTGCTTAA 480
 Db 421 AGACCTATGCTGCGACCCCGATCGGCGCTGTACCTTACTCTGTTGCTGCTTAA 480
 QY 481 CCTAGCAAGGAC 492
 Db 481 CCTAGCAAGGAC 492

RESULT 4
 AAX09010 standard; cDNA; 1272 BP.
 ID AAX09010
 XX
 AC AAX09010:
 XX
 DT 14-JUN-1999 (first entry)
 XX
 DE Brn-3a polynucleotide.
 XX
 KW Brn-3a; Bcl-2; neurone; neuronal cells; apoptosis; cell death; CNS;
 KW PNS; central nervous system; parasympathetic nervous system;
 KW development; injury; neurotrophic factor; nerve growth factor; NGF;
 KW ciliary neurotrophic factor; CNTF; brain-derived neurotrophic factor;
 KW BDNF; neurotrophin; NT-3; NT-4; NT-5; neurodegenerative disease;
 KW familial dysautonomia; infantile muscular dystrophy;
 KW Parkinson's disease; Alzheimer's disease; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key 1.1.3972
 FH CDS location/Qualifiers
 FT //tag= a
 FT /product= Brn-3a-polypeptide
 XX

RESULT	6
AA017186/c	AA017186 standard; DNA; 125401 BP.
XX	
XX	
AC	AA017186;
XX	
DT	29-NOV-2001 (first entry)
XX	
DE	Streptomyces noursei nystatin PKS gene cluster DNA.
XX	
KW	Polypeptide synthase; PKS; macrolide; nystatin; PKS gene cluster;
KW	antifungal; antibiotic; ds.
XX	
OS	Streptomyces noursei.
XX	
Key	Location/Qualifiers
FH	6337..34771
FT	/tag- a
FT	/product- "NysI complete protein"
FT	34792..51099
FT	/tag- b
FT	/product- "NysJ protein"
FT	51155..57355
FT	/tag- c
FT	/product- "NysK protein"
FT	57503..58687
FT	/tag- d
FT	/product- "NysL protein"
FT	complement (58786..58980)
FT	/tag- e
FT	/product- "NysM protein"
FT	/note- "CDS does not include start codon"
FT	/tag- f
FT	/product- "NysN protein"
FT	/note- "CDS does not include start codon"
FT	complement (60238..61296)
FT	/tag- g
FT	/product- "NysO complete protein"
FT	120628..121308
FT	/tag- h
FT	/product- "NysP (long) protein"
XX	
XX	MO200159126-A2.
PN	
XX	
PD	16-AUG-2001.
XX	
PE	08-FEB-2001; 2001MO-GH0509.
XX	
PR	08-FEB-2000; 2000GB-0002840.
PR	10-APR-2000; 2000GB-0008786.
PR	14-APR-2000; 2000GB-0009387.
XX	
PA	(UNIV-) UNIV NORGES TEKNISK NATURVITTENSKAPSELIGE.
PA	(SWE-) SINTER STIPELSEN IND TEK FORSK.
PA	(ALPH-) ALPHARMA AS.
PA	(SIN-) SINVENT AS.
PA	(DIE/) DIEGLEWSKA H.
PA	(ZOT/) ZOTCHEV S B.
PA	(SEKU/) SEKUROVA O N.
PA	(FJAE/) FJAEVRIK E.
PA	(BRAU/) BRAUTASET T.
PA	(STRO/) STROM A R.
XX	
XX	
PI	Zotchev SB, Sekurova ON, Fjaervik E, Brautaset T, Strom AR;
PI	Valla S, Ellingsen TE, Sletta H, Gulliksen O;
XX	
DR	WPI: 2001-557614/62.
DR	P-PSDS: AA010143, AA010144, AA010145, AA010146, AA010147, AA010148,
DR	AA010149, AA010150.
XX	
XX	New nystatin polypeptide synthase polynucleotides and polypeptides,

xx	Claim 1; Page 188-254; 266pp; English.
xx	The present invention relates to the cloning and sequencing of the gene
xx	cluster encoding a modular type I polyketide synthase (PKS) enzyme
xx	involved in the biosynthesis of the macrolide antibiotic nystatin.
xx	The nystatin PKS is useful as antifungal antibiotics. The present
xx	sequence is a Streptomyces noursei nystatin PKS gene cluster DNA.
SQ	Sequence 125401 BP; 15664 A; 49692 C; 42871 G; 17174 T; 0 other;
	Query Match 9.5%; Score 46.8; DB 22; Length 125401;
	Best Local Similarity 52.6%; Pred. No. 0.26;
	Matches 102; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
OY	23 CGCGCCGCAACTCTTCGGGATATCCTTCGTCCCTCCTCACCGGCACCCCATGGTAGC 82
Dd	44196 CGGTACCGCGGTGTGCGCCGTGCTCTGTCTGTAGCGCCGACGACGAGACACGCGGTGCG 44137
OY	83 GGCCAGCATCGACCCTTGCCTGTGGAAAGCTGTACATGCTGATGCGCGCGCTGGTGCGGC 142
Dd	44136 GCGCGGTGTGTGTGGCGCCGTTCTGTGTAGCAGGCGCGGCGAAGCGGCCGCTCGTCCGTG 44077
OY	143 GCGCGCGGTCTTCGCCCTGTCTCGCGGATGCGCGGTGCGCGGCTTGAGCTGTCCGCGCGCGC 202
Dd	44076 GCGCGGAGCATCTCACCCCCGCGGCGGTCCGCGGCGAGGCGGCGAGAGAGCTGCGGCACCA 44017
OY	203 GCGCGATGAGGCGG 216
Dd	44016 GGGGTGTGGTGTGCG 44003
RESULT 7	
AHAH41550/C	
ID	AHAH41550 standard; DNA; 2765 BP.
XX AC	AHAH41550;
XX DT	14-SEP-2001 (first entry)
XX DE	Human Rlt1 gamma nucleotide sequence.
XX KW	Mouse; human; combined DNA/RNA molecule; Rlt1; tumour suppressor;
KW	2-3 type zinc finger structure; cancer; diagnosis; carcinogenesis;
XX	gene therapy; ds.
OS	Homo sapiens.
XX	
FH Key	Location/Qualifiers
FT CDS	11..2116
FT	/tag=a
FT	/product= "Rlt1 gamma"
PX	WO200132859-A1.
PN	
PD	10-MAY-2001.
PX	
PF	14-JUL-2000; 2000WO-JP04765.
PR	
XX	29-OCT-1999; 99JP-0310420.
PA	(MOCH) MOCHIDA PHARM CO LTD.
PI	Komlunani R;
DR	WPI: 2001-316438/33.
DR	P-PSDB: AAB99341.
PT	New zinc finger protein and gene encoding it for detecting and
PT	diagnosing cancer, estimating the risk of carcinogenesis, and for gene
PT	therapy -
XX	

PT	therapy
XX	
PS	Claim 38; Page 73-76; 119pp; Japanese.
CC	
CC	The present invention describes a combined DNA/RNA molecule designated
CC	Ritl1, which has a 2-3 type zinc finger structure and tumour suppressor
CC	activity. Ritl1 has cytosstatic activity and can be used in gene therapy.
CC	Genomic or cDNA encoding Ritl1 can be used in the detection and diagnosis
CC	of cancer, and the estimation of the risk of carcinogenesis. Ritl1 and
CC	its partial peptides are also used to detect and diagnose cancer, and
CC	estimate the risk of carcinogenesis. The present sequence encodes
CC	human Ritl1 alpha', from the present invention.
XX	
SQ	Sequence 3332 BP; 637 A; 1113 C; 1036 G; 534 T; 12 other;
Query Match	9.3%; Score 46; DB 22; Length 3332;
Best Local Similarity	50.5%; Pred. No. 0.25;
Matches 112; Conservative	0; Mismatches 110; Indels 0; Gaps 0;
OY	32 CTCCTGCGGGATATTCCTGTCTCCTCTCCACCGCGCACCCCCAGTGAAGGCCCAACTGC 91
Db	1639 CTCCAGTAGACAGCT 1580
OY	92 GCGCCCTGCTGGGAAAGCTGTACATGCTGATCGGCGGGCGTGCGGTCGCCGCCGCCGGCTC 151
Db	1579 GCCCAGCGACGGGTTCCCTCTGTGGTGGCGGAATGACCTGTGCGCCCTTGAGGCCCTTC 1520
OY	152 TTCCGCTGTGTCGGCGGTGCGCGGTCCGATCGGCGCTTTGGCGTCCGGCGCGCGCATGA 211
Db	1519 GCCCGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1460
OY	212 GGGCGGCACTGTGGTGTGATTCACGACCACCTGAGGGTCAACAAT 253
Db	1459 GCGGCGGCGCACGACGACGCGGCTTGTGATGTGCGCTTCAAT 1418
RESULT 10	
AAH41505/c	
ID	AAH41505 standard; DNA: 3344 BP.
XX	
AC	AAH41505;
XX	
DT	14-SEP-2001 (first entry)
XX	
DE	Human Ritl1 encoding DNA SEQ ID NO:3.
XX	
KW	Mouse: human; combined DNA/RNA molecule; Ritl1; tumour suppressor;
KW	2-3 type zinc finger structure; cancer; diagnosis; carcinogenesis;
KW	gene therapy; ds.
XX	
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	CDS 11..2695
FT	/*tag= a
FT	/product= "Ritl1"
XX	
PN	WO200132859-A1.
PD	10-MAY-2001.
XX	
PP	14-JUL-2000; 2000WD-JP04765.
PR	29-OCT-1999; 99JP-0310420.
XX	
PA	(MOCH) MOCHIDA PHARM CO LTD.
PI	Komlnami R;
XX	
DR	WI: 2001-316438/33.
XX	
DR	P-PDB: AAB99335.
XX	
PT	New zinc finger protein and gene encoding it for detecting and

```

PT diagnosing cancer, estimating the risk of carcinogenesis, and for gene
PT therapy
PS Claim 8; Page 103-108; 119pp; Japanese.
XX
XX
XX The present invention describes a combined DNA/RNA molecule designated
CC Rtl1, which has a 2-3 type zinc finger structure and tumour suppressor
CC activity. Rtl1 has cytosstatic activity and can be used in gene therapy.
CC Genomic or cDNA encoding Rtl1 can be used in the detection and diagnosis
CC of cancer, and the estimation of the risk of carcinogenesis. Rtl1 and
CC its partial peptides are also used to detect and diagnose cancer, and
CC estimate the risk of carcinogenesis. The present sequence encodes
CC human Rtl1.
XX
XX
SQ Sequence 3344 BP; 639 A; 1115 C; 1040 G; 538 T; 12 other:
XX
Query Match 9.3%; Score 46; DB 22; Length 3344;
Best Local Similarity 50.5%; Pred. No. 0.25;
Matches 112; Conservative 0; Mismatches 110; Indels 0; Gaps 0;
QY 32 CTCGCGGAGATATCCCTGCTCCTCCTCCACCGGACACCCCATGATGAGCGCCAGCTC 91
Db 1651 CTCGAGTAGAGAGCTCCTCCTCCTCCTCCTCCGTCGTCGTCCTCCTCCTCCGCGTCGTC 1592
QY 92 GCGCCCTCCCTGGGAAAGCTGATCATGCTGATCGCGCGGCTGTCGCGCGCCGCGCTC 151
Db 1591 GCCCGACGACGGGTGCTCTCTGCTGGGGAAGTCACCGTCGGCGCCCTTGAGGCCCTC 1532
QY 152 TTCGCGCTGCTGCGGGGCGGCGGCTGTCGCGGCGCTTGCGGCTGCGCGCGCGCGATGA 211
Db 1531 GCCCGCCAGCTGCTGCTGTCGCGGCTCGGGGAGCTGCGCGCGCGGAGAGCCCTGCTCGGA 1472
QY 212 GGGCGGACCTGGGTGTGATCAGCCACCATGAGGCTCAACAT 253
Db 1471 GCGCGCGGCGCAGCGAGCGGCGCTTGATGATGCGCTTTCAT 1430
RESULT 11
AAA95215
ID AAA95215 standard; cDNA; 530 BP.
XX
AC AAA95215;
XX
DT 12-JAN-2001 (first entry)
XX
DE Human OCT-T1 partial coding sequence.
XX
KW GC rich nucleic acid; DNA amplification; zwitterion;
KW base pairing disruption; polymerase chain reaction; fragile X syndrome;
KW spinobulbar muscular atrophy; myotonic dystrophy; Huntington's disease;
XX Huntington's chorea; spinocerebellar ataxia type 1; ss.
XX
OS Homo sapiens.
XX
XX US6114150-A.
XX
XX 05-SEP-2000.
XX
XX 02-DEC-1996; 96US-0758662.
XX
XX 29-NOV-1995; 95US-0564653.
XX
XX (UYYA ) UNIV YALE.
XX
XX Weissman SM, Baskaran N;
XX
XX WPI; 2000-586478/55.
XX
XX
XX Uniform amplification of heterogeneous mixture of nucleic acid
XX templates of varying G+C content, comprises amplifying the reaction
XX mixture in the presence of betaine and dimethyl sulphoxide
XX
XX Example 1; column 13-14; 15pp; English.
XX

```

DR P-PSDB;AAW68541.
XX
PT New diagnosis of Leukaemia - by detecting genes for tumour antigen
FT rejection precursors or corresponding proteins
XX
PS Claim 1; Pages 63-67; 88pp; English.

The present sequence represents the nucleotide sequence of the
CC SIX DP2-64 (Oct-TI) gene. The protein is a tumour rejection
CC antigen precursor (TRAP). The specification describes the treatment of
CC disorders which characterised by expression of a Leukemia-associated
CC nucleic acid such as TRM. The products are used for in vivo or in vitro
CC screening for leukemia, lymphoma or other cancers by usual
CC hybridisation/amplification or immunoassay methods. TRAPS, when
CC processed to antigens or complexed with HLA (human lymphocyte antigen)
CC molecules, or nucleic acid encoding them, are useful in vaccines for
CC treating leukaemia.

XX
SO Sequence 4524 BP; 1177 A; 1090 C; 1180 G; 1077 T; 0 other:

Query Match 9.3%; Score 45.6; DB 20; Length 4524;
Best Local Similarity 54.3%; Pred. No. 0.32;
Matches 113; Conservative 0; Mismatches 94; Indels 1; Gaps 1

OY 22 TCGGGCGGAACCTCCGCGGGAGATCCTGTCTCCTCCTCACCGGCACCCTCATGTAG 81
Db 523 TGGACCAATCATCTCCGCGCTCGCGTATGAGCGCGCGCGCGCGCGCGCG 582
OY 82 CGGCGAGCTCGCGCGCTCGGAAAGCTGACATGCTG-ATGCGGCGCTGCGGCCG 140
Db 583 CGGCGGCG 642
OY 141 GCGCGCGGCTTCGCGCTGCTGCGCGGTGCGGTCCTGCGCGCTTGCGCGTCCGCGCG 200
Db 643 GCGGCG 702
OY 201 GCGCGCGATGAGGCGCGCACCTGGGTGG 228
Db 703 GCGGCGCGCGCGCGCGCGCTCCTGCGCGG 730

RESULT 13
AAQ83793/C
ID AAQ83793 standard; cDNA; 1146 BP.
XX
XX AAQ83793;
XX
XX 05-SEP-1995 (first entry)

S. avermitilis BCKDH EI-alpha subunit.
XX
XX S. avermitilis BCKDH EI-alpha subunit.
XX
XX Branched-chain alpha-keto-acid-dehydrogenase; BCKDH; bkd gene;
KW avermectin: antibiotic; acaricide; anthelmintic; insecticide;
KW nematocide; pesticide; ds.
XX
XX Streptomyces avermitilis.
OS
XX
FH Key Location/Qualifiers
FT 1..1146
FT CDS /**tag= a
XX
XX MO504150-A.
XX
XX PN
XX PD 09-FEB-1995.
XX
XX PE 30-MAY-1994; 94MO-IB00127.
XX
XX PR 30-JUL-1993; 93US-0100518.
XX
XX PA (PRIZ) PFIZER INC.
XX
XX PI Denoya CD;
XX

DR	WPI: 1995-082233/11.
DR	P-PSDB: AAR69623.
XX	
PT	Streptomyces branched-chain alpha-keto-acid dehydrogenase DNA
PT	used to improve prodn. of natural avermectin cpds and to produce
PT	novel avermectin cpds
XX	
PS	Claim 5; Page 42; 65pp: English.
XX	
CC	Novel branched-chain alpha-keto-acid-dehydrogenase (BCKDH) genes
CC	(bkd) from Streptomyces avermectilis were cloned using PCR and
CC	homology probing. DNAs encoding the BCKDH E1-alpha, E1-beta, and
CC	N-terminal and internal portions of the E2 subunit (AA083793-96)
CC	were obtained. A genomic sequence including the E1-alpha, E1-beta
CC	and E2 (partial) bkd ORFs was also isolated (AA083797). Manipulation
CC	of bkd genes allows the enhanced production of natural or novel
CC	avermectins.
XX	
XX	
SQ	Sequence 1146 BP; 163 A; 465 C; 392 G; 126 T; 0 other;

[illegible]

RESULT 14	AAQ83797/C	AAQ83797 standard; cDNA; 2728 BP.
ID	AAQ83797	standard; cDNA; 2728 BP.
XX	AAQ83797;	
XX	AC	
XX	DT	05-SEP-1995 (first entry)
XX	XX	
DE	S. avermitilis bkd gene region.	
XX		
XX	Branched-chain alpha-keto-acid-dehydrogenase; BCKDH; bkd gene	
KW	avermectin; antidiabetic; acaricide; anthelmintic; insecticide;	
KW	nematocide; pesticide; ds.	
XX		
OS	Streptomyces avermitilis.	
XX		
FH	Key	Location/Qualifiers
FT	RBS	390..395
FT		/*tag= a
FT	CDS	403..1548
FT		/*tag= b
FT		/product= E1-alpha subunit
FT	RBS	1607..1614
FT		/*tag= c
FT	CDS	1622..2626
FT		/*tag= d
FT		/product= E1-beta subunit
FT	RBS	2610..2615
FT		/*tag= e
FT	CDS	2626..2728
FT		/*tag= f
FT		/product= E2 N-terminal sequence

XX WO9504150-A.
 PN
 XX
 PD 09-FEB-1995.
 XX
 PF 30-MAY-1994; 94WO-IB00127.
 XX
 PR 30-JUL-1993; 93US-0100518.
 XX
 PA (PFIZ) PFIZER INC.
 XX
 PI Denoya CD;
 DR WPI: 1995-082233/11.
 DR P-PSDB; AAR69623; AAR69624; AAR69625.
 XX
 XX Streptomyces branched-chain alpha-keto-acid dehydrogenase DNA -
 PT used to improve prodn. of natural avermectin cpds and to produce
 PT novel avermectin cpds
 XX
 PS Claim 5; Page 44-45; 65pp; English.

Query Match	9.0%;	Score 44.2;	DB 16;	Length 2728;
Best Local Similarity	52.4%;	Pred. No. 0.62;		
Matches 97;	Conservative 0;	Mismatches 88;	Indels 0;	Gaps 0;

QY	52	CCTCTCTCCACAGCGGACCCCATGATGAGCGGCCAGCTCGCGCCCTGCTGGAAAGCT	111
Db	1204	CGCGCTCCACAGAGCTCGCGCCCTCCCTCCGCGCGCTGGGCCACAGCGCTGAGAGA	1145
QY	112	GTACATGCTGATTCGGCGGCGTGGGCGGCGGCGCGGCTTCCGCGCTGCTCGCGGGTGC	171
Db	1144	CCTGTTGCACAGGCGCGCGCGCTGTCCTGACACAGCGCGCCGCGCATCCCGTAGCCGA	1085
QY	172	CGGTCGCGGCGCCCTTGGCGTCCGCGGCGGCGCGCGGATGAGAGCGCGCACCTGGGTGTGA	231
Db	1084	CGGCGCTGTGGGCCACGACGAGGCGGCGGCTGTCTTGGCGAGCGGAGCAGATGGCGA	1025
QY	232	TCCAG	236
Db	1024	AGCCG	1020

RESULT 15
ABV23346
ID ABV23346 standard; cDNA: 1388 BP.
XX
AC ABV23346;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 23337.
XX
KW Human; prostate cancer; cytosolic; carcinogen; pharmacodynamic marker;
XX
RV pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN W0200160860-A2.
XX
PD 23-AUG-2001.
XX

```

PE 20-FEB-2001; 2001MO-US05171.
PR 17-FEB-2000; 2000US-183319P.
PR 16-MAR-2000; 2000US-183862P.
PR 25-MAY-2000; 2000US-207454P.
PR 09-JUN-2000; 2000US-211314P.
PR 18-JUL-2000; 2000US-219007P.
PR 13-DEC-2000; 2000US-255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer -
XX
PS Claim 1; Page 4234-4235; 11750dp; English.
CC
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for:
CC (a) assessing whether a patient is afflicted with prostate cancer;
CC (b) monitoring the progression of prostate cancer in a patient;
CC (c) assessing the efficacy of a test compound to inhibit prostate
CC cancer in a patient;
CC (d) assessing the efficacy of a therapy for inhibiting prostate cancer
CC in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound;
CC (g) determining whether prostate cancer has metastasized in a patient;
CC (h) assessing the aggressiveness or indolence of prostate cancer in a
CC patient;
CC (I) is also useful as a pharmacodynamic or pharmacogenomic marker.
XX
SQ Sequence 1388 BP; 205 A; 445 C; 377 G; 338 T; 23 other:
XX
Query Match      8.9%; Score 44; DB 23; Length 1388;
Best Local Similarity 52.7%; Pred. No. 0.63;
Matches 118; Conservative 0; Mismatches 105; Indels 1; Gaps 1
QY 8 GATCATCTTCAACATCGGCGGCACACTCCTGGGGATATCCTGTCTCCTCCAGCAG 67
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1148 GTTTCCTCCGCGATGCGTGCTCTTCATCTGCTCACTCTGCGGCTGTTCCGCCGTT 1207
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 68 CACCCCATTGGTAGGGGCGCACGTGCGGCCCTCTCTGGGAACAAGTGAATGCTGATCGC 127
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1208 CTTCATCATAGTCATGATGAGCGGCGCGGCGCTCGTCCGCGGCGCCCTTGATTAAGCCATG 1267
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 128 GCGCTCGGTGCGCGGCGCGGCGGTCTTTCGCGCTGCT-CGCGGAGTGCCTGCGGCGCT 186
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1268 AGCGGCGGCGGCGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGACAGCGGCGAG 1327
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 187 TGGCGTCCGCGGCGCGCGCATGAGAGGCGGACCTGTGGGTGTG 230
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1328 CGGCGAGCGGCGGTGCGCGCGAACAAGTACGTGACGAGGCGGCGCG 1371

```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 25, 2002, 22:44:43 ; Search time 2088 Seconds
(without alignments)
3816.178 Million cell updates/sec

Title: US-09-548-449-3

Perfect score: 492
Sequence: 1 actcgcgacatctacc.....gccttaacctagaagac 492

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 100%
Listing first 45 summaries

Database :

EST :

- 1: em_estba:*
- 2: em_estlum:*
- 3: em_estlin:*
- 4: em_estnu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_hnc:*
- 9: gb_estl:*
- 10: gb_est2:*
- 11: gb_hnc:*
- 12: gb_est3:*
- 13: gb_est4:*
- 14: gb_est5:*
- 15: em_estlum:*
- 16: em_estom:*
- 17: gb_gss:*
- 18: em_gss_hum:*
- 19: em_gss_inv:*
- 20: em_gss_pln:*
- 21: em_gss_vrt:*
- 22: em_gss_fun:*
- 23: em_gss_mam:*
- 24: em_gss_mus:*
- 25: em_gss_other:*
- 26: em_gss_pro:*
- 27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59.4	12.1	925	17	CNS0091P
2	53.6	10.9	1101	17	AL108460 Drosophila
3	51	10.4	12	BG487612	EM1_65_B1
4	49.8	10.1	523	12	BF266741 HV_C8A001
5	48.8	9.9	610	10	BE426427 WHE0335_H
6	48.2	9.8	764	9	AL547921 AL547921

Result No.	Score	Query Match	Length	DB ID	Description
7	47.6	9.8	332	10	BE597437
8	47.6	9.7	130	12	BF485240
9	47.6	9.7	392	12	BG356596
10	47.6	9.7	529	10	AM678427
11	47.6	9.7	536	10	AM747229
12	47.6	9.7	587	10	AM680372
13	47.6	9.7	617	10	AM678141
14	47.4	9.6	803	17	AG162326
15	47.2	9.6	697	10	AM155496
16	47.2	9.6	920	10	AM155435
17	47.2	9.6	935	17	CNS006XK
18	47.2	9.6	497	12	BG366604
19	46.8	9.5	617	10	BE195746
20	46.8	9.5	840	17	AG043467
21	46.6	9.5	932	17	CNS00720
22	46.4	9.4	367	13	BM318362
23	46.4	9.4	571	10	AM672287
24	46.4	9.4	844	17	CNS0052P
25	46.2	9.4	631	9	AL508579
26	46	9.3	925	17	CNS0091P
27	46	9.3	1008	14	BO708747
28	46	9.3	1071	17	CNS00EMV
29	45.6	9.3	485	10	BE599458
30	45.6	9.3	488	13	BM376745
31	45.6	9.3	535	13	BM329587
32	45.6	9.3	558	14	BQ467551
33	45.6	9.3	561	12	BF619248
34	45.6	9.3	587	13	BM329864
35	45.6	9.3	632	12	BG708176
36	45.6	9.3	841	12	BF255640
37	45.6	9.3	848	12	BG310135
38	45.6	9.3	951	12	BG343611
39	45.6	9.3	993	12	BG343243
40	45.6	9.3	1219	14	BM803425
41	45.6	9.3	1511	14	BO900625
42	45.4	9.2	632	10	BE414408
43	45.4	9.2	736	14	BQ839017
44	45.2	9.2	338	13	BI306090
45	45.2	9.2	497	10	AM564083

ALIGNMENTS

RESULT 1
CNS0091P
LOCUS
DEFINITION
CNS0091P 925 bp DNA linear GSS 03-JUN-1999
BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.

ACCESSION
AL053013
VERSION
AL053013.1 GI:4934461
KEYWORDS
GSS.

SOURCE
Drosophila melanogaster.

ORGANISM
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequence :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).

The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazuhiro Oosawa and
Aaron Mammoler in Pletier de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial


```
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone_lib="Embryo 1 (EM1)"
/note="Organ: Embryos germinated for 24 hr; Vector:
pbluescript II from lambda zap II; Site 1: XhoI; Site 2:
EcoRI; The library was made from poly-A RNA in the cloning
vector lambda zap II. Clones to be sequenced were
prepared by mass excision."
BASE COUNT      82 a      179 c      218 g      74 t
ORIGIN

Query Match      10.4%; Score 51; DB 12; Length 553;
Best Local Similarity 47.9%; Pred. No. 0.14;
Matches 147; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

QY 2 CTCGCGATATCTTACACATCGGCGCAACTCTGCGGATATCTCTCTCTC 61
    || || || || || || || || || || || || || || || || ||
Db 382 CTGAACACACGCGCGCTGATCTGACACGCGCTCGGACACTCGCGCTG 323
    || || || || || || || || || || || || || || || || ||
QY 62 CACCGGACACCCCATGATGAGCGGCGGCGGCGGCGGCGGCGGCGGCG 121
    || || || || || || || || || || || || || || || || ||
Db 322 CTCTGCTGGCAACCCACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 263
    || || || || || || || || || || || || || || || || ||
QY 122 ATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 181
    || || || || || || || || || || || || || || || || ||
Db 262 GCAAGTACGCTGTGAGGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 203
    || || || || || || || || || || || || || || || || ||
QY 182 GGCCTGACGCTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 241
    || || || || || || || || || || || || || || || || ||
Db 202 CACGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 143
    || || || || || || || || || || || || || || || || ||
QY 242 GAGGTCACATTCCTCACTCCGCGGAAATGATTCCTGATGATGCGCCACG 301
    || || || || || || || || || || || || || || || || ||
Db 142 GCGGAGCCACCGCCAGGCTCCAGAGGCGACGCGGCGGCGGCGGCGGCGG 83
    || || || || || || || || || || || || || || || || ||
QY 302 CCGCGG 308
    || || || || || || || || || || || || || || || || ||
Db 82 CACCGCG 76

RESULT 4
BF266741/c 523 bp mRNA linear EST 23-OCT-2001
LOCUS      BF266741
DEFINITION HV_Cea0015024f Hordeum vulgare seedling green leaf EST library
            HVCDA0004 (Blumeria challenged) Hordeum vulgare cDNA clone
            HV_Cea0015024f, mRNA sequence.
ACCESSION  BF266741
VERSION     BF266741.2 GI:13262953
KEYWORDS   EST.
SOURCE     Hordeum vulgare.
ORGANISM   Hordeum vulgare.
REFERENCE  Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae
            ; Triticeae; Hordeum.
            1 (bases 1 to 523)
            Wing, R., Close, T.J., Kleinhofs, A., Wise, R., Wei, F., Begum, D.,
            Fritsch, D., Yu, Y., Henry, D., Palmer, M., Rambo, T., Simmons, J., Choi
            , D.W., Fenton, R.D., Oates, R. and Main, D.
            Development of a genetically and physically anchored EST resource
            for barley genomics: Blumeria infected incompatible (Mla13)
            seedling leaf cDNA library
            Unpublished (2001)
            On Nov 17, 2000 this sequence version replaced gi:11197736.
            Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel.: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Total hg bases = 276
            Seq primer: AATTAACTTCATCAAGCG
            High quality sequence stop: 400.

JOURNAL
COMMENT
```

```
FEATURES
Source
Location/Qualifiers
1..523
/organism="Hordeum vulgare"
/cultivar="C116155 (Mla13)"
/db_xref="taxon:4513"
/clone="HV_Cea0015024f"
/clone_lib="Hordeum vulgare seedling green leaf EST
library HVCDA0004 (Blumeria challenged)"
/tissue_type="seedling green leaf"
/lab_host="TJC121"
/note="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI;
C.I. 16155 (Mla13) plants were greenhouse grown in the R
Wise lab at Iowa State University, Ames, IA; 7 day old
green seedlings were challenged with isolate A27 (AvrMla13
) of Blumeria graminis f. sp. hordei, and leaves were
harvested 20 and 24 hr post-inoculation and snap frozen;
uninoculated leaves were harvested 20 hr post-inoculation
(Wei, Wise). In the TJ Close lab at the University of
California, Riverside, total RNA was prepared from each
sample pool, equal quantities of all three RNA pools were
combined, poly(A) RNA was purified from the mixture, one
cDNA library was made, and 1 million pfu were in vivo
excluded to give pBluescript SK(-) cDNA phagemids (Choi,
Close). Phagemids were plated and picked at the Clemson
University Genomics Institute (CUGI) (Begum, Palmer,
Fritsch, Atkins and Wing). Plasmid DNA preparations, DNA
sequencing and sequence analysis were performed at CUGI
(Wing, Yu, Fritsch, Henry, Simmons, Oates, Rambo, Main).
The sequence has been trimmed to remove vector sequence
and contains a minimum of 100 bases of phred value 20 or
above. For more details on library preparation and
sequence analysis see
http://www.genome.clemson.edu/projects/barley. To order
this clone see http://www.genome.clemson.edu/orders
see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
Genetically and physically anchored EST resources for
barley genomics. Barley Genetics Newsletter 31:29-30.
(http://wheat.pw.usda.gov/g9pages/bgn/31/cover.html)"
BASE COUNT      89 a      156 c      210 g      67 t
ORIGIN

Query Match      10.1%; Score 49.8; DB 12; Length 523;
Best Local Similarity 50.9%; Pred. No. 0.26;
Matches 117; Conservative 0; Mismatches 113; Indels 0; Gaps 0;

QY 4 CGCGATCATCTTCACATCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 63
    || || || || || || || || || || || || || || || || ||
Db 411 CGAAGTTAAACAGCACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 352
    || || || || || || || || || || || || || || || || ||
QY 64 CCGGACCCCGCATGCTAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 123
    || || || || || || || || || || || || || || || || ||
Db 351 CCGGCTCGGCTTGGCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 292
    || || || || || || || || || || || || || || || || ||
QY 124 CCGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 183
    || || || || || || || || || || || || || || || || ||
Db 291 NTACATCATCTTCGCTGGAATTCGTGACGCGGCGGCGGCGGCGGCGGCG 232
    || || || || || || || || || || || || || || || || ||
QY 184 CTTGCGGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 233
    || || || || || || || || || || || || || || || || ||
Db 231 GAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 182
    || || || || || || || || || || || || || || || || ||

RESULT 5
BE426427/c 610 bp mRNA linear EST 24-JUN-2000
LOCUS      BE426427
DEFINITION WHE0335_H02_P03S wheat unstressed seedling shoot cDNA library
            BE426427
            Triticum aestivum cDNA clone WHE0335_H02_P03, mRNA sequence.
ACCESSION  BE426427
VERSION     BE426427.1 GI:9424270
KEYWORDS   EST.
SOURCE     bread wheat.
ORGANISM   Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
; Triticeae; Triticum.
1 (bases 1 to 610)
Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D., Han
, P.S., Hsia, C.C., Kang, Y., Jazo, G.R., Miller, R., Rausch, C.J.,
Seaton, C.L. and Tong, J.C.
The structure and function of the expressed portion of the wheat
genomes
Unpublished (2000)
Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific
West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105595773
Fax: 5105595818
Email: andersn@wpr.usda.gov
Sequence have been trimmed to remove vector sequence and low
quality sequence with phred score less than 20
Seq primer: Stralagene 5K primer.

FEATURES	Location/Qualifiers
source	1. .610

```

/organism="Trifolium aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE0335_H02_P03"
/clone_id="Wheat unstressed seedling shoot cDNA library"
/tissue_type="Etiolated shoot"
/dev_stage="Five day old seedling"
/lab_host="E. coli SOLr"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid:
Site_1: EcoRI; Site_2: XhoI; Seeds were surface-sterilized,
germinated and grown aseptically in the dark at room
temperature on filter paper with water, nystatin and
cefotaxime in covered crystallization dishes. Shoots were
harvested. The tissue, total RNA, and poly(A) RNA were
prepared, a cDNA library was made, and the cDNA clones
were in vivo excised to give plasmid phagemids in the
TJ Close lab (Choi, Close, Fenton) at the University of
California, Riverside. Plasmid DNA preparations and DNA
sequencing were performed in the OD Anderson lab (all
other authors)."

```

Query Match	9.9%	Score 48.8	DB 10	Length 610
Best Local Similarity	51.9%	Pred. No. 0.46		
Matches 110	Conservative	0	Mismatches 102	Indels 0
Gaps				
QY	32	CTCTCGGGGATATCTGTCCTCTCTCTCACGGGACACCCCATGATGAGCGGCAGCTC	91	
Db	542	CTCTTCGACGAGATGCGCGTCCCGACAGCGCCGACGCGCGGTGAGTGCCTTGCC	483	
QY	92	GGCGCCGCGGCGGGAAGCTGTACATGCTATGCGCGGCGTCCGTCGGCGCGCGGCTC	151	
Db	482	GGCATCTGCTCCGCGCTGCTATGCTGTGCGTCACTCCGTCGGGCGGTGCGCGCGGA	423	
QY	152	TTCCGCGTGTCTCGGCGGCGGTCCGCGCGGCGTTCGCGTCCGCGCGCGCGATGA	211	
Db	422	GGTGGCGTGGCGGCGCGTCCGCGCGGCGTCCGCGCGGCGGCGGCGGCGGCGGCGG	363	
QY	212	GGGCGGCGACCTGGGTGTGATCAGCCACTGA	243	
Db	362	AGACGAAAGCGAAGAGAGACGAAAGCCGACGA	331	
RESULT 6				
LOCUS	AL547921/c			
DEFINITION	AL547921 prime, mRNA sequence.	764 bp	mRNA	linear
ACCESSION	AL547921			
VERSION	AL547921.1			
KEYWORDS	EST.			
				EST 16-FEB-2001
				clone CS0D1033YH07 3

SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 764)
TITLE	Li, W. B., Gruber, C., Jesssee, J. and Polayes, D.
JOURNAL	Full-length cdna libraries and normalization
COMMENT	Unpublished (2001)
FEATURES	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seq@genoscope.cns.fr, Web : www.genoscope.cns.fr. Location/Qualifiers

FEATURES	location/qualifiers
source	1. .764

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0D1033YH07"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-Oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com url : http://fulllength.invitrogen.com"
```

[illegible]

RESULT	7
LOCUS	BE597437/c
DEFINITION	332 bp mRNA linear EST 18-AUG-2000 P11.69_E06_bL_A002 pathogen induced 1 (P11) Sorghum bicolor cDNA,
ACCESSION	BE597437
VERSION	mRNA sequence.
KEYWORDS	BE597437
SOURCE	BE597437.1 GI:9852510
ORGANISM	EST. Sorghum.
REFERENCE	Sorghum bicolor Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoidae; Andropogoneae; Sorghum. 1 (bases 1 to 332) Cordonnier-Pratt,M.-M., Gingle,A., Dean,R., Sudman,M. and Pratt, .U.H.
TITLE	An EST database from Sorghum: pathogen-induced plants
JOURNAL	Unpublished (2000)
COMMENT	Contact: Cordonnier-Pratt MM Laboratory for Genomics and Bioinformatics

The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmp@plantuga.edu

Seq primer: JEN REV
High quality sequence stop: 288
POLYA=NO.

FEATURES

Source

Location/Qualifiers

```

1..332
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone_lib="pathogen induced 1 (PII)"
/note="Organ: Anthracnose infected leaves from
two-week-old sorghum plants 48 hr after inoculation;
Vector: pBluescript II from lambda Zap II; Site.1: XhoI;
Site.2: EcoRI; Two-week-old sorghum plants (BRX 623
cultivar) were infected with pathogen (isolate FR4421 of
collotrichum graminicola, which is a sorghum isolate).
RNA was prepared from infected leaves harvested from 45
seedlings 48 hours after inoculation. Note: young
seedlings (2 weeks old) exhibit juvenile resistant
reaction, which is an incompatible interaction. As they
grow older (4 weeks or older), plants resume susceptibility
to anthracnose disease. The library was made from poly-A
RNA in the cloning vector lambda Zap II. Clones to be
sequenced were prepared by mass excision. WARNING: While
most or all ESTs are expected to derive from the host
plant, no effort was made to eliminate ESTs deriving from
the pathogen."

```

BASE COUNT	43 a	107 c	140 g	42 f
ORIGIN				

Query Match	9.88;	Score 48;	DB 10;	Length 332;
Best Local Similarity	51.98;	Pred. No. 0.59;		

QY	14	CTTACCACTGCGGCGCACTCTGGGGGATATCCCTGCTCTCTCCACCGGAGCCCC	73
Db	324	CTTGCCCCCGCCGATCTCTTACGACGCGCCACGAGCGCTTCGTCGCGGCCAAGCST	265
QY	74	CATGGTAGCGGCCACGCTCGCGCCCTGCTGGGAAAGCTGTACATGCTGTATCGCGCGCTC	133
Db	264	CTTGGCGCGCCGACGATGCGCCCGGACAGACGAGAACTGTCTGTGGCCGGCGGCGAC	205
QY	134	GGTGCAGGCGGCGCGGGGTCTTCCGCTCTCGGCGGTGCGCGTCCGTGCGGCTTGGCGCTC	193
Db	204	GGCGTCGACGACGACAGCGGCTCGGCGACGATCTCTGCTGGCGCGGACGACGACGAC	145
QY	194	CGCGGCGGCGCGCATGTAGGGGGGAGC	221
Db	144	CTCGCGGTGCGCGCCCTTCGGCGGAGGCC	117

RESULT 8	
BF485240	
LOCUS	190 bp mRNA linear EST 06-DEC-2000
DEFINITION	WHEI1790.A11.B22S wheat pre-anthesis spike cDNA library Trilicium aestivum cDNA clone WHEI1790.A11.B22. mRNA sequence.

ACCESSION	BF485240
VERSION	BF485240.1
GI	GI:11568541

KEYWORDS	EST.
SOURCE	bread wheat.
ORGANISM	Triticum aestivum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticeae; Triticum.

TITLE	The structure and function of the expressed portion of the wheat genomes - Pre-anthesis spike cDNA library
JOURNAL	unpublished (2000)
COMMENT	Contact: Olin Anderson

FEATURES

Source

Location/Qualifiers

"organism="Triticum aestivum"
"cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE1790.A11.B2"
/clone_lib="wheat pre-anthesis spike cdna library"
/tissue_type="Spike before anthesis"
/lab_stage="Adult plant"
/db_host="E. coli SOLR"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site_1: EcoRI; Site_2: XhoI; Plants were grown in the
greenhouse. Whole spike with awns trimmed, white, green
and yellow anther were collected and total RNA, and
poly(A) RNA were prepared, a cDNA library was made, and
the cDNA clones were in vivo excised to give phagescript
phagemids in the '93 Close lab (Choi, Close, Fenton) at
the University of California, Riverside. Plasmid DNA
preparations and DNA sequencing were performed. In the OD
Anderson lab (all other authors)."

BASE COUNT	26 a	67 c	70 g	27 t
ORIGIN				

Query Match	9.7%	Score 47.6;	DB 12;	Length 190;
Best Local Similarity	62.7%	Pred. No. 0.62;		
Matches 74;	Conservative 0;	Mismatches 44;	Indels 0;	Gaps 0

[illegible]

RESULT 9					
EG356596/c					
LOCUS	EG356596	392 bp	mRNA	linear	EST 06-MAR-2001
DEFINITION	EM1_24_E04.g1_A002 Embryo 1 (EM1) Sorghum bicolor cDNA, mRNA				

sequence.	
BG356596	ACCESSION
BG356596.1	VERSION
GI:13238582	

KEYWORDS	E.
SOURCE	SC

ORGANISM
Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.

REFERENCE
1 (bases 1 to 392)
Reid, S.P., Cordonier-Pratt, M.-M., Gingle, A. and Pratt, L.H.

TITLE	An EST database from sorghum: developing embryos
JOURNAL	Unpublished (2000)
COMMENT	Contact: Cordonnier-Pratt MM

BASE COUNT	ORIGIN
97 a	144 c 178 g 117 t

Query Match	9.7%	Score 47.6	DB 10	Length 536
Best Local Similarity	51.9%	Pred. No. 0.83		
Matches 107	Conservative 0	Mismatches 99	Indels 0	Gaps 0

[illegible]

Db 223 GCGGAGAGTAAATTGGATTCATTGGCCGTCGATAGCGGGGGGGCCGGATCGAGCTGGGCGCT 164

Dy 163 GCGGCGATGCGGCTCTGTCGGGCTTTGGCGTCCGCGGCGGCGCGGATGAGGGCGGCACT 222

Db 163 CGACGAGCGCCACGGGTCCTTTCCACAGGTGAACTTGGCCGCGCGGCCACGACTTGGGCGGCG 104

Qy 223 GGGTGGTGATCCAGCCACTGAAGGTC 248
| | | | | | | | | |
Db 103 GCGTCCGATACCCAGCCGCGGGGC 78

RESULT 12			
AM680372/c	587 bp	mRNA	linear EST 19-JUL-2000
LOCUS	AM680372		
DEFINITION	WS1_52_A04.q1_A002	water-stressed 1 (WS1)	sorghum bicolor cdna, mRNA sequence.

VERSION	AW680372.1	GI:75541611
KEYWORDS	EST.	
SOURCE	sorghum.	

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
clade; Panicoideae; Andropogoneae; Sorghum.

AUTHORS: Cordonnier Pratt, M.-M., Gingle, A., Marsala, C., Sudman, M. and Pratt, L.H.

TITLE	An EST database from Sorghum: water-stressed plants
JOURNAL	Unpublished (2000)
COMMENT	Contact: Cordonnier-Pratt MM

The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
E-mail: TAC@UGA.EDU

Tel: 706 542 1860
Fax: 706 583 0210
Email: empratt@uga.edu

Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.

High quality sequence start: 68
High quality sequence stop: 587
POLYA=Yes.

FEATURES	Location/Qualifiers
source	1. .587

```
/organism="Sorghum bicolor"  
/db_xref="taxon:4558"
```

```

/ab_xref="cdx01:4536"
/clone_lib="Water-stressed 1 (WS1)"
/notes="Organ: Mix of 5-week old na

```

water was withheld; Vector: Lambda Zap; Site_1: XhoI; Site_2: EcoRI. The library was made from poly-A RNA in the

BASE COUNT	112 a	153 c	207 g	114 t	1 others
	prepared by mass excision.				

BASE C
ORIGIN

Query Match	9.7%	Score 47.6;	DB 10;	Length 587;
Best Local Similarity	51.9%;	Pred. No. 0.85;		
Matches 107; Conservative	0;	Mismatches 99;	Indels 0;	Gaps 0;

350 TAGCCACCTTCACCTCGCCGACGACCACTTCCTTAGCCACCTCCTCCTCCTCCTCGG 291

290 GCCCGACGACTGTTGGATCATCTTGGCTGCATGCGCGCGCGGGCCGGGTGCAGCTGGCCCT 231
163 CGGCGGTGCCGTCCTGCGGCCCTTGCGCTCCGCGCGCGCGCATGAGGCGGACACT 224

223 GGGTGGTATCCAGCCACTGAGGGTC 248
 170 GCCTGCCGATACCCACGCGGGGGG 145

RESULT 13	
W678141/c	
LOCUS	617 bp
AM678141	mRNA
DEFINITION	linear EST 19-JUL-2000
MS1_13_B09.g1_A002	water-stressed 1 (WS1) Sorghum bicolor cDNA,

VERSION	AW678141.1	GI:7551863
KEYWORDS	EST.	
SOURCE	sorghum.	

Chlorophyta; Charophyta; Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum.

REFERENCE
1 (bases 1 to 617)
Cordonnier-Pratt, M.-M., Gingle, A., Marsala, C., Sudman, M. and Pratt, L.H.

TITLE An EST database from Sorghum: water-stressed plants
JOURNAL Unpublished (2000)
COMMENT Contact: Cordonnier-Pratt MM

Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
E-mail: zhu@arches.uga.edu

Tel: 706 542 1860
Fax: 706 583 0210

Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.

High quality sequence start: 87
High quality sequence stop: 617
POLYA=Yes.

FEATURES	Location/Qualifiers
source	1. .617

```
1:101/  
/organism="Sorghum bicolor"  
/db_xref="taxon:4558"
```

```

/clone_1lb="Water-stressed 1 (WS1)"
/note="Organ: Mix Of 5-week old plants on days 7 & 8 after

```

water was withheld; Vector: lambda Zap; Site_1: XhoI; Site_2: EcoRI; The library was made from poly-A RNA in the cloning vector lambda Zap II. Clones to be sequenced were

	prepared by mass excision.			
BASE COUNT	117 a	162 c	215 g	123 t
ORIGIN				

Query Match	9.78;	Score 47.6;	DB 10;	Length 617;

Best Local Similarity 51.9%; Pred. No. 0.86;
Matches 107; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

43 TATTCCTGTCCTCCCTCCACCGGCAACCCCATGTGTAAGCGGCGAAGCTCGCGCCCTGCT 10

364 TAGCCACCTCCACCTGGCGAGACACCTCCCTAGCACCTCTCTCTCTCTCCG 307

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 25, 2002, 20:22:07 ; Search time 56 Seconds
(without alignments)
2694.374 Million cell updates/sec

Title: US-09-548-449-3

Perfect score: 492
Sequence: 1 actcgcgcagcatcctacc.....gccttaacctagcaagac 492

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata1/1na/5A_COMB.seq:*
2: /cgn2_6/ptodata1/1na/5B_COMB.seq:*
3: /cgn2_6/ptodata1/1na/6A_COMB.seq:*
4: /cgn2_6/ptodata1/1na/6B_COMB.seq:*
5: /cgn2_6/ptodata1/1na/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata1/1na/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45.6	9.3	530	3	US-08-758-662-4
2	45.6	9.3	4524	2	US-08-845-898-7
3	45.6	9.3	4524	3	US-09-206-537-7
4	45.6	9.3	4524	4	US-09-430-854-7
5	44.2	9.0	1146	1	US-08-482-385A-1
6	44.2	9.0	2728	1	US-08-482-385A-5
7	44.2	8.9	4257	2	US-08-690-473-1
8	44.2	8.9	4257	4	US-09-259-821A-1
9	44.2	8.9	4257	4	US-08-843-659-1
10	44.2	8.9	12001	1	US-08-458-568A-11
11	43.8	8.9	68750	3	US-09-335-409-1
12	43.8	8.9	68750	4	US-09-568-102-1
13	43.8	8.9	68750	4	US-09-567-969-1
14	43.8	8.9	68750	4	US-09-568-480-1
15	43.8	8.9	68750	4	US-09-568-486-1
16	43.8	8.9	68750	4	US-09-568-472-1
17	43.8	8.9	68750	4	US-09-567-899-1
18	43.8	8.9	71989	4	US-09-443-501A-2
19	42.8	8.7	367	4	US-09-072-596-249
20	42.8	8.7	1028	4	US-08-118-700-1
21	42.8	8.7	1028	4	US-08-458-745-1
22	42.4	8.6	2846	4	US-09-613-182-5
23	42.4	8.6	44377	2	US-08-804-227C-7
24	42.4	8.6	44377	2	US-08-804-198-1
25	42.2	8.6	1926	4	US-09-249-585A-4
26	42.2	8.6	1931	2	US-09-130-114-2
27	41.2	8.4	4403765	4	US-09-103-840A-2

C 28	41	8.3	861	2	US-08-924-759-15	Sequence 15, Appl
C 29	41	8.3	861	3	US-09-248-335-15	Sequence 15, Appl
C 30	41	8.3	1279	3	US-09-248-335-25	Sequence 25, Appl
C 31	40.8	8.3	2214	3	US-08-864-038A-2	Sequence 1, Appl
C 32	40.8	8.3	2889	3	US-08-537-002A-4	Sequence 4, Appl
C 33	40.8	8.3	2889	3	US-08-863-010-4	Sequence 4, Appl
C 34	40.8	8.3	3331	3	US-09-024-429-4	Sequence 4, Appl
C 35	40.8	8.3	3331	3	US-08-864-038A-2	Sequence 2, Appl
C 36	40.8	8.3	3331	3	US-08-864-038A-4	Sequence 4, Appl
C 37	40.8	8.3	3600	1	US-08-537-002A-5	Sequence 5, Appl
C 38	40.8	8.3	3600	4	US-08-863-010-5	Sequence 5, Appl
C 39	40.8	8.3	3600	4	US-09-024-429-5	Sequence 5, Appl
C 40	40.6	8.3	2185	1	US-08-173-508-3	Sequence 3, Appl
C 41	40.6	8.3	2185	2	US-08-265-310-3	Sequence 3, Appl
C 42	40.6	8.3	2185	2	US-08-951-742-3	Sequence 3, Appl
C 43	40.4	8.2	16442	3	US-08-781-891-208	Sequence 208, App
C 44	40.2	8.2	304	4	US-09-056-556-166	Sequence 166, App
C 45	40.2	8.2	304	4	US-09-072-596-161	Sequence 161, App

ALIGNMENTS

```
RESULT 1
US-08-758-662-4
; Sequence 4, Application US/08758662
; Patent No. 6114150
; GENERAL INFORMATION:
; APPLICANT: Weisman, Sherman
; APPLICANT: Baskaran, Namadev
; TITLE OF INVENTION: AMPLIFICATION OF NUCLEIC ACIDS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 701 Fifth Avenue, 6300 Columbia Center
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/758,662
; FILING DATE: 29-Nov-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6114150tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 390036.402C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: Linear
; US-08-758-662-4
;
Query Match          9.3% Score 45.6; DB 3; Length 530;
Best Local Similarity 54.3%; Pred. No. 0.057;
Matches 113; Conservative 0; Mismatches 94; Indels 1; Gaps 1;
QY 22 TCGGCGGCACTCGGCGGATATCTGCTCTCCACCGGACCCGCAATGCTAG 81
Db 264 TGGACGACATCTCTCGGCGGCTCGGCTCGGCTCGGCGGCGGCGGCGGCGG 323
QY 82 CGGCGACATCGGCGGCTCGGCGGAAACGCTTACATGCTG-ATGCGGCGGCTGCGGCG 140
||||| 11 11 111 11 11 111 111 111 111 111 111 111 111 111
```

[illegible]

RESULT 2
fig 00 01

```

1      Sequence 7, Application US/08845998
2      Patient No. 5879892
3
4      GENERAL INFORMATION:
5      APPLICANT: Van Baren, Nicolas
6      APPLICANT: Coule, Pierre G.
7      APPLICANT: De Smet, Charles
8      APPLICANT: Lucas, Sophie
9      APPLICANT: Boon, Thierry
10     TITLE OF INVENTION: LEUKEMIA ASSOCIATED GENES
11     NUMBER OF SEQUENCES: 16
12
13     CORRESPONDENCE ADDRESS:
14     ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
15     STREET: 600 Atlantic Avenue
16     CITY: Boston
17     STATE: MA
18     COUNTRY: US
19     ZIP: 02210
20
21     COMPUTER READABLE FORM:
22     MEDIUM TYPE: Floppy disk
23     COMPUTER: IBM PC compatible
24     OPERATING SYSTEM: PC-DOS/MS-DOS
25     SOFTWARE: Patent Release #1.0, Version #1.25
26     CURRENT APPLICATION DATA:
27     APPLICATION NUMBER: US/08/845,998
28     FILING DATE:
29
30     CLASSIFICATION: 435
31
32     ATTORNEY/AGENT INFORMATION:
33     NAME: Van Amsterdam, John R.
34     REGISTRATION NUMBER: 40,212
35     REFERENCE/DOCKET NUMBER: L0461/7008
36     TELECOMMUNICATION INFORMATION:
37     TELEPHONE: (617)720-3500
38     TELEFAX: (617)720-2441
39
40     INFORMATION FOR SEQ ID NO: 7:
41     SEQUENCE CHARACTERISTICS:
42     LENGTH: 4524 base pairs
43     TYPE: nucleic acid
44     STRANDEDNESS: double
45     TOPOLOGY: linear
46     MOLECULE TYPE: cDNA
47     HYPOTHEetical: NO
48     ANTI-SENSE: NO
49
50     FEATURE:
51     NAME/KEY: CDS
52     LOCATION: 174..1433
53
54     US-08-845-998-7

```

Query Match	9.38;	Score 45.6;	DB 2;	Length 4524;
Best Local Similarity	54.38;	Pred. No. 0.072;		
Matches 113;	Conservative 0;	Mismatches 94;	Indels 1;	Gaps 1

QY	22	TGGGCGGCAATCTCGGCGGATATCTGTCCTCTCCACCGGACCCCATGTAG	81
Db	523	TGGACCAATCTCTCGCGCGCTGCTCGCGCTCATGGCGCGCGGCGGCGCGG	582
QY	82	CGGCGAGTGGCGGCGGCGGGAAGCTGTAACTGTG-ATGGGGGGGCTGCTGCGG	140
Db	583	CGCGCGGCGGCGGCGGCGGCCACACAGCGCCCGGCGGCGGCTGTGGCGGCCCGGCGGCGG	642
QY	141	CGGCGCGGCTTCCGCTGCTCGGCGGATGCGGCTGCCGAGGCTTGGCGCTCCGCGGCG	200

Db	Qy	Db
643	GCAGCCGATGAGGGCGGACCTGGGTG 228	702
703	GCAGCCGATGAGGGCGGACCTGGGTG 730	

RESULT 3

Sequence 7, Application US/09206537
 Patent No. 6130052
 GENERAL INFORMATION:
 APPLICANT: Van Baren, Nicolas
 APPLICANT: Coulle, Pierre G.
 APPLICANT: De Smet, Charles
 APPLICANT: Lucas, Sophie
 APPLICANT: Boon, Thierry
 TITLE OF INVENTION: LEUKEMIA ASSOCIATED GENES
 NUMBER OF SEQUENCES: 16
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
 STREET: 600 Atlantic Avenue
 CITY: Boston
 STATE: MA
 COUNTRY: US
 ZIP: 02210
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/206,537
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/845,998
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Van Amsterdam, John R.
 REGISTRATION NUMBER: 40,212
 REFERENCE/DOCKET NUMBER: L0461/7008
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617)720-3500
 TELEFAX: (617)720-2441
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 4524 base pairs
 TYPE: nucleic acid
 TOPOLOGY: double
 STRANDEDNESS: linear
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: NO
 ANTI-SENSE: NO
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 174..1433
 US-09-206-537-7

Query Match	9.38;	Score 45.6;	DB 3;	Length 4524;
Best Local Similarity	54.38;	Pred. No. 0.072;		
Matches 113; Conservative	0;	Mismatches 94;	Indels 1;	Gaps 1

[illegible]

Patent No. 5728561
GENERAL INFORMATION:
APPLICANT: DENOVA,, CLAUDIO D.
TITLE OF INVENTION: GENES ENCODING BRANCHED CHAIN ALPHA KETOACID DEHYDROGENASE FROM STREPTOMYCES SPECIES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: PETER C. RICHARDSON
STREET: 235 EAST 42ND STREET, 20TH FLOOR
CITY: NEW YORK
STATE: NEW YORK
COUNTRY: U.S.A
ZIP: 10017-5755

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPILER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/M5-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,385A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SHEYKA,, ROBERT F.
REGISTRATION NUMBER: 31,304
REFERENCE/DOCKET NUMBER: PC8346C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-573-1189
TELEFAX: 212-573-1939

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2728 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA

US-08-482-385-A-5

Query Match 9.0%, Score 44.2; DB 1; Length 2728;
Best Local Similarity 52.4%; Pred.No.0.14;

Matches 97; Conservative 0; Mismatches 88; Indels 0; Gaps 0.

Dy 52 CCTGCTCCGCCAGCGGCACCCTGGATGTGAACGCGCACACTTCGCGCCCTGCTTGGAAAGCT 111
||| ||| | | | | | | | | | | |
Db 1204 CGCGCTTCACGAGCGCTTGCGCCCCCTCCCAGCGCGCATGCAAGCGGTGCTGAGA 1145

Qy 112 GTACATCTGATACGCGGCGCTGCGTGCACGCGCGCGGTCTTCCGCTCTGCGCGGTGC 171
||| | | | | | | | | | | | | |
Db 1144 CCTGCTGCACGCGCGCCCGCGCTGTGTTGCCGTGCACAGCGCGCGCCGCACTCCGTAGCCA 1085

Qy 172 CGGTCGTCGCGCCTTGGCGCTCGCGCGCGCGCGCGCATGAGGCGGCGCACCTGGGTGTGA 231
|||| | | | | | | | | | | | | |
Db 1084 CGGCGCTTGTGGCGCACGACAGGCGGCGCGGTCTGCTTGGCGAGCGGAGCAGAGATGGCGA 1025

Yr 232 TCCAG 236
|| |
DB 1024 AGCGG 1020

RESULT 7
US-08-690-473-1/C
Sequence 1, Application US/08690473
Patent No. 5876923
GENERAL INFORMATION:
Applicant: Leopardl, Rosario
ApplcAnt: Roizman, Bernard
Title Of Invention: HERPES SIMPLEX VIRUS ICP4 AS AN
INHIBITOR OF APOPTOSIS
Number Of Sequences: 2
Correspondence Address:
Addressee: Arnold, White & Durkee
Street: P.O. Box 4433
City: Houston

```

: STATE: Texas
: COUNTRY: USA
: ZIP: 77210
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/690.473
: FILING DATE: 26-JUL-1996
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Highlander, Steven L.
: REGISTRATION NUMBER: 37,642
: REFERENCE/DOCKET NUMBER: ARCD:239
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 512/418-3000
: TELEFAX: 512/474-7577
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4257 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
: US-08-690-473-1
:
: Query Match 8.9%; Score 44; DB 2; Length 4257;
: Best Local Similarity 50.5%; Pred. No. 0.16;
: Matches 107; Conservative 0; Mismatches 105; Indels 0; Gaps 0;
:
: Oy 19 CCATCGGGCCGCAACCTCCGCGGATATCTCTCTCTCTCCATCCGACCGGACCCCAATGG 78
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
: Db 613 CCAGTGTGTCGGCGGACGAGCGGTCGCCGCGCTCTCGCGCGGTGCGGTGCGCGCGG 554
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
: Oy 79 TAGCGGCACGCTCGGCGCCCTCGCTGGGAAGCTGTAACTGATCGGCGCGGCTCGGTC 138
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
: Db 553 GGGCCCTCCCGTCCCGCGCGGCGGTCTGAGAGTCGTGGGGGTGTCGCGGTGTCGTCGG 494
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
: Oy 139 CGGCGCGCGGCTTCGCGCTGCTCGGCGGTGCGGATCGGCTGAGCTTGGCGTCCGCG 198
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
: Db 493 GGTCTCCCGCGCTCTCCGCTCGTCCGCGGCGCCACCGGAGGCGCCCGCTCGTGGCGGT 434
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
: Oy 199 CGGCGCGGATGAGGCGCGGACCTTGGGTGTCG 230
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
: Db 433 CTGGGCTCGGGGTGGGCGGCGCCCGCTCGGTG 402
: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
:
: RESULT 8
: US-09-259-821A-1/c
: Sequence 1, Application US/09259821A
: Patent No. 6210926
: GENERAL INFORMATION:
: APPLICANT: LEOPARDI, ROSARIO
: APPLICANT: ROIZMAN, BERNARD
: TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICp4 IS AN INHIBITOR OF APOPTOSIS
: FILE REFERENCE: ARCD:317
: CURRENT APPLICATION NUMBER: US/09/259.821A
: CURRENT FILING DATE: 1999-03-01
: PRIOR APPLICATION NUMBER: 08/690.473
: PRIOR FILING DATE: 1996-07-26
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: Patentin Ver. 2.1
: SEQ ID NO 1
: LENGTH: 4257
: TYPE: DNA
: ORGANISM: HERPES VIRUS, TYPE 1
: US-09-259-821A-1
:
: Query Match 8.9%; Score 44; DB 4; Length 4257;
: Best Local Similarity 50.5%; Pred. No. 0.16;
: Matches 107; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

```

[illegible]

RESULT 9
US-08-843-659-1/c
; Sequence 1, Application US/08843659
Default US 6010103

GENERAL INFORMATION:
APPLICANT: Leopard, Roasario
TITLE OF INVENTION: Roisman, Bernard
TITLE OF INVENTION: HERPES SIMPLEX VIRUS US3 AND ICP4 AS
INHIBITORS OF APOPTOSIS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P. O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/843, 659
FILING DATE: Concurrently Herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: ARSB:519
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4257 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

Query Match	8.9%	Score 44	DB 4	Length 4257
Similarity	50.5%	Pred. No. 0.16		
Matches 107	Conservative	0	Mismatches 105	Indels 0
			Gaps	0
QY	19	CCATCGGCCGCAAC	CTCGCGGAAATCCTCTCTCTCCCTCCACACGGACCCCATGG	78
Db	613	CCAGCTGTCGCGCGG	CAGAGCGCTCCCGCGCTCTCGCGGGCGTCGTCCCGCGCGG	554
QY	79	TAGCGGCAAGCTCGCGCCCTCGTGGGAAAGCTGTACATGTCGATCGCGGCGCTCGGTC	138	
Db	553	GGGCGCTCCCGGCGCGCGGCGCTCGTCGAGAGTCTGNGGGGTGTCGCGGTCGTGGTCGG	494	
QY	139	CGGGGCGCGGCTTCCTCCGCTCTCGCGCGTCGCCGTCCTGTCGCGGCTTTGGGCTCCGCG	198	
Db	493	GCTGTTCCTCCCGGCTCTCTCGTCTCCGCGCCACCCGAGAGGCCCTCCGCTCTCGCGGT	434	

```

QY      199  CGCGCGCCATGAGCGCGCACCTGGGTGTG  230
          | | | | | | | | | | | | | | | |
Db      433  CTGGGCTGGGGTGGGCGGCGCCGTCGGTG  402

```

RESULT 10
US-08-458-568A-11
; Sequence 11, Application US/08458568A

GENERAL INFORMATION:
APPLICANT: Schaffer, Priscilla A.
APPLICANT: Yeh Lily
TITLE OF INVENTION: Compositions and Methods for Treatment of Herpesvirus
TITLE OF INVENTION: Infections
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5821339r1s
STREET: One Liberty Place, 46th floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,568A
FILING DATE: 02-JUNE-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/065,146
FILING DATE: 05-MAY-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Leary Ph.D., Kathya R.
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: DPCI-0029
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 12001 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Herpes simplex virus
STRAIN: Herpes Simplex Virus type 1
US-08-458-568A-11

[illegible]

cy 199 CGCGCGCGCATGAGGGCGGCACCTGGGTGTC 230
 | | | | | | | | | | | | | | | | | | | |
 Db 6240 CTGGGCTCGGGGTGGGCGGCGCCGTCGGTG 6271

RESULT 11

Sequence 1, Application US/09335409
 Patent No. 6121029
 GENERAL INFORMATION:
 APPLICANT: Schupp, Thomas
 APPLICANT: Ligon, James
 APPLICANT: Molnar, Istvan
 APPLICANT: Zitzke, Ross
 APPLICANT: Cyr, Devon
 TITLE OF INVENTION: GENES FOR THE
 FILE REFERENCE: 4-30582A
 CURRENT APPLICATION NUMBER: US/09/
 CURRENT FILING DATE: 1999-06-17
 NUMBER OF SEQ. ID NOS.: 30
 SOFTWARE: Patentln Ver. 2.0
 SEQ. ID NO. 1
 LENGTH: 66750
 TYPE: DNA
 ORGANISM: Sorangium cellulosum
 US-09-335-409-1

Query Match	8.9%	Score	43.8	DB	3	Length	68750
Best Local Similarity	52.5%	Pred.	NO	0.24			
Matches	96	Conservative	0	Mismatches	87	Indels	0
						Gaps	0

QY 14 CTTCACACGAGGCGCAATCTCTGGGGGATATACCGTCCTCCACACAGGCGACGCC 73
 Db 63905 CTACCGAGGGAGATTGTCTCTCGCTCGCTCGCGCTCGCGCTCTCTCTGACCGG 63966
 QY 74 CATGTATAGCGGACAGCTCGCGGCTCGCTGGGAAAGCTATCATCTGATTCGGGCGGTC 133
 Db 63965 GCGGCTTCGGCGAGCTTCGCGGGGCGCTGCGCCAGGCCGAGGTCTCGGCGAGCTCTTC 64024
 QY 134 GGTATCCGCGGCGCGGGTCTTTCGCCCTGCTCGGCGGATGCCGATCGTACGAGCTTTGCGTC 193
 Db 64025 GGGGCGCTGGTCTGTGGGCCCGCTCTGTCGTCCGCGCGCTCTCTTGGGTTTCATTCAGTTC 64084

QY	194	CGC	196
		11	
Db	64085	CTC	64087

RESULT 12

```

Sequence, Application US/09568102
Patient No. 6346404
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James
APPLICANT: Molnar, Istvan
APPLICANT: Zirkle, Ross
APPLICANT: Cyr, Devon
APPLICANT: Goelach, Joein
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/568, 102
CURRENT FILING DATE: 2000-05-10
PRIOR APPLICATION NUMBER: 09/335,409
PRIOR FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 68750
TYPE: DNA
ORGANISM: Sorangium cellulosum
US-09-568-102-1

```

Query Match	8.98;	Score 43.8;	DB 4;	Length 68750;
Best Local Similarity	52.58;	Pred. No. 0.24;		
Matches 96;	Conservative 0;	Mismatches 87;	Indels 0;	Gaps 0;

QY	14	CTGACGCAATGGGCGCAACTCCCTGGGGGATATCTGTCGTCCTCTCCACCGGCGACCC	73
Db	63905	CTACACGAGAGCGCAGSTTCTGCTCTCCGCTCGTCGCGCTCGCGCTCTCTCTGACCGCG	63964
QY	74	CATGTAGCGGCACACTCGCGCCCTGCGGAAAGCTGTACATGCTGATCGGCGCGTC	133
Db	63965	CGCGCCTTCGCGACCTCGCGCGGCGCGGCCACGACCGAGAGCTCGGCGAGCTCTTC	64024
QY	134	GGTCCCGCGCGCGCGGCTCTCCGCGCTGCTCGGCGGGTGGCGGTCGCGGACCTTGGCGTC	193
Db	64025	GGCGGCGTGTGTCGTGGGCGCGTCGCTGTCGCGGCGGCTGCTCCGTTTCATGAGATC	64084
QY	194	CGC	196
Db	64085	CTTC	64087

RESULT 13

```

: Sequence 1, Application US/09567969
: Patent No. 635457
:
: GENERAL INFORMATION:
: APPLICANT: Schupp, Thomas
: APPLICANT: Lignon, James
: APPLICANT: Molnar, Istvan
: APPLICANT: Zirkle, Ross
: APPLICANT: Cyr, Devon
: APPLICANT: Goelach, Joern
: TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
: FILE REFERENCE: 4-30582A
: CURRENT APPLICATION NUMBER: US/09/567, 969
: CURRENT FILING DATE: 2000-05-10
: PRIOR APPLICATION NUMBER: 09/335,409
: PRIOR FILING DATE: 1999-06-17
: NUMBER OF SEQ ID NOS: 30
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1
: LENGTH: 68750
: TYPE: DNA
: ORGANISM: Sorangium cellulosum
: US-09-567-969-1

```

Query Match	Score	DB	Length
8.98;	43.8;	4;	68750;

Matches	96;	Conservative	0;	Mismatches	87;	Indels	0;	Gaps	0;
---------	-----	--------------	----	------------	-----	--------	----	------	----

QY	14	CTACACATTCGGCCGCACTCCCTGGGGATATCTCTGCTCTCTCTCTACACGGCACCC	73
Db	63905	CTACACGACGGGCAAGTTCTGTCTCTGCTCGTCCGCTCGCGCTGTCTCTCTGACCGG	63964
QY	74	CATGTACCGGCGACATCTCGCGCCCTGCCTCGGGAAAGCTGTACATGCTATCGGCGCGTC	133
Db	63965	CGCGGCTTCGGGAGACTGGCGGGGGGGTGGGCCAGGCCGAGTCTGGCGCAGCTTTC	64024
QY	134	GGTGGCGGCGCGCGGCTTCGCGCTCGCTCGGGGCTGCGGCTGCGGCGGCTTGGCGTC	193
Db	64025	GCGCGCTGTGTCGTGGGCCCGCTCGTCTCGTGGCGCGGCTGCTCGGTTCATGAGTTC	64084
QY	194	CGC	196
Db	64085	CTC	64087

RESULT 14

US-09568-480-1
; Sequence 1, Application US/09568480
; Patent NO. 6355458
; GENERAL INFORMATION:

APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James
APPLICANT: Molnar, Istvan
APPLICANT: Zirkle, Ross
APPLICANT: Cyr, Devon
APPLICANT: Goelzsch, Joern
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/568,480
CURRENT FILING DATE: 2000-05-10
PRIOR APPLICATION NUMBER: 09/335,409
PRIOR FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 68750
TYPE: DNA
ORGANISM: Sorangium cellulosum
US-09-568-480-1

Query Match 8.9%; Score 43.8; DB 4; Length 68750;
Best Local Similarity 52.5%; Pred. No. 0.24;
Matches 96; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 14 CTTCACCATGCGCCGCACTCTGCGGGATATCTCTCTCTCTCAACGGACCC 73
DB 63905 CTCACCGAGGCGAGGTCTGCTGCTGCGCCCTGCTCTGACCCGCG 63964
QY 74 CATGTAGCGGCGCAGCTCGCGCCCTGCGTGGAAAGCTGTACATCGATCGGGCGCTC 133
DB 63965 CGGCGCTTGGCGGAGCTCGCGCGCGGCTCGCCAGCCGAGGTCTCGGCGAGCTTTC 64024
QY 134 GGTGCGGCGCGCGCTTCTGCGCTGCTGCGGGGCGGCGCGCTCGGCTTGGCGTC 193
DB 64025 GCGGCGGTGGTGTGGCGCGCTCGCTGCTGCGGCGGCTCGCTCGGTTCATCGAGTC 64084
QY 194 CGC 196
DB 64085 CTC 64087

RESULT 15
US-09-568-486-1
Sequence 1, Application US/09568486
Patent No. 6355459
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James
APPLICANT: Molnar, Istvan
APPLICANT: Zirkle, Ross
APPLICANT: Cyr, Devon
APPLICANT: Goelzsch, Joern
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/568,486
CURRENT FILING DATE: 2000-05-10
PRIOR APPLICATION NUMBER: 09/335,409
PRIOR FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 68750
TYPE: DNA
ORGANISM: Sorangium cellulosum
US-09-568-486-1

Query Match 8.9%; Score 43.8; DB 4; Length 68750;
Best Local Similarity 52.5%; Pred. No. 0.24;
Matches 96; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

QY 14 CTTCACCATGCGCGCACTCTGCGGGATATCTCTCTCTCTCAACGGACCC 73
DB 63905 CTCACCGAGGCGAGGTCTGCTGCTGCGCCCTGCTCTGACCCGCG 63964

QY 74 CATGTAGCGGCGCAGCTCGCGCCCTGCGTGGAAAGCTGTACATCGATCGGGCGCTC 133
DB 63965 CGGCGCTTGGCGGAGCTCGCGCGCGGCTCGCCAGCCGAGGTCTCGGCGAGCTTTC 64024
QY 134 GGTGCGGCGCGCGGCTTCTGCGCTGCTGCGGGGCGGCGCGCTCGGCTTGGCGTC 193
DB 64025 GCGGCGGTGGTGTGGCGCGCTCGCTGCTGCGGCGGCTCGCTCGGTTCATCGAGTC 64084
QY 194 CGC 196
DB 64085 CTC 64087

Search completed: December 26, 2002, 06:11:24
Job time : 396 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 25, 2002, 22:45:58 : Search time 236 Seconds

(without alignments)
846.895 Million cell updates/sec

Title: US-09-548-449-3

Perfect score: 492
Sequence: 1 actcgcgcatcattcacc.....gccttaactagaagac 492

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 363474 seqs, 203117208 residues

Total number of hits satisfying chosen parameters: 726948

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PC7_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PC7US_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45.6	9.3	530	US-10-073-353-4	Sequence 4, Appl1
2	44.4	8.9	4257	US-09-825-288A-1	Sequence 1, Appl1
3	43.8	8.9	68750	US-10-014-717-1	Sequence 1, Appl1
4	43.4	8.8	1131	US-09-712-363-93	Sequence 93, Appl1
5	43.4	8.8	1305	US-09-815-242-7730	Sequence 7730, Ap
6	41.8	8.5	1360	US-09-867-550-1927	Sequence 1927, Ap
7	41.1	8.3	1665	US-09-815-242-7935	Sequence 7935, Ap
8	40.2	8.2	1428	US-09-921-332-5	Sequence 5, Appl1
9	40.2	8.2	1428	US-09-921-330-5	Sequence 5, Appl1
10	40.2	8.2	1428	US-09-921-329-5	Sequence 5, Appl1
11	40.2	8.2	2481	US-09-894-998-35	Sequence 35, Appl1
12	39.6	8.0	4446	US-09-815-242-7983	Sequence 7983, Ap
13	39.4	8.0	550	US-09-791-171-63	Sequence 63, Appl1
14	39.4	8.0	2561	US-09-976-740-48	Sequence 48, Appl1
15	39.4	8.0	2561	US-10-023-529-48	Sequence 48, Appl1
16	39.4	8.0	2561	US-10-023-523-48	Sequence 48, Appl1
17	39.4	8.0	3032	US-09-954-043-1	Sequence 1, Appl1
18	39.4	8.0	3152	US-09-880-107-3431	Sequence 3431, Ap
19	39.2	8.0	1771	US-09-991-436-7	Sequence 7, Appl1

C 20	39.2	8.0	1771	10	US-09-874-923-7	Sequence 7, Appl1
C 21	39	7.9	684	10	US-09-969-708-52	Sequence 52, Appl1
C 22	39	7.9	684	10	US-09-880-107-1604	Sequence 1604, Ap
C 23	38.8	7.9	4039	10	US-09-205-448-7	Sequence 7, Appl1
C 24	38.6	7.8	1436	10	US-09-788-345-11	Sequence 11, Appl1
C 25	38.6	7.8	1518	10	US-09-815-242-7829	Sequence 7829, Ap
C 26	38.6	7.8	7419	10	US-09-815-242-4009	Sequence 4009, Ap
C 27	38.4	7.8	424	10	US-09-960-352-14599	Sequence 14599, A
C 28	38.4	7.8	429	9	US-09-854-133-413	Sequence 413, App
C 29	38.4	7.8	429	10	US-09-738-973-113	Sequence 413, App
C 30	38.4	7.8	1406	10	US-09-745-763-18	Sequence 18, Appl1
C 31	38.4	7.8	3624	10	US-09-927-112-3	Sequence 3, Appl1
C 32	38.4	7.8	4635	10	US-09-927-112-3	Sequence 3, Appl1
C 33	38.4	7.8	4635	10	US-10-001-873-22	Sequence 22, Appl1
C 34	38	7.7	1508	9	US-09-931-457A-61	Sequence 61, Appl1
C 35	38	7.7	2162	10	US-09-799-875-10	Sequence 10, Appl1
C 36	38	7.7	2586	9	US-09-905-291A-235	Sequence 235, App
C 37	38	7.7	2586	9	US-10-063-547-7	Sequence 7, Appl1
C 38	38	7.7	2586	9	US-09-902-853-235	Sequence 235, App
C 39	38	7.7	2586	10	US-09-909-320-235	Sequence 235, App
C 40	38	7.7	2586	10	US-09-909-088B-235	Sequence 235, App
C 41	38	7.7	2586	12	US-10-006-667-7	Sequence 7, Appl1
C 42	37.8	7.7	8509	10	US-09-826-205-1	Sequence 1, Appl1
C 43	37.4	7.6	371	10	US-09-867-550-299	Sequence 299, App
C 44	37.4	7.6	3042	9	US-09-712-363-68	Sequence 68, Appl1
C 45	37.4	7.6	3203	10	US-09-880-107-2164	Sequence 2164, Ap

ALIGNMENTS

RESULT 1
US-10-073-353-4
Sequence 4, Application US/10073353
Patent No. US2002016858A1

GENERAL INFORMATION:

APPLICANT: Sherman M. Weissman

NAME: Namdev Baskaran

TITLE OF INVENTION: Amplification of Nucleic Acids

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Morgan, Lewis & Bockius LLP

STREET: 1111 Pennsylvania Ave., NW

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20004

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0. Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/073,353

FILING DATE: 03-May-2002

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/758,662

FILING DATE: 1996-12-07

APPLICATION NUMBER: US 08/758,662 (CPA)

FILING DATE: 1999-02-17

ATTORNEY/AGENT INFORMATION:

NAME: Michael S. Tuscan

REGISTRATION NUMBER: 43,210

REFERENCE/DOCKET NUMBER: 44921-5007-03-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-739-3001

TELEFAX: 202-739-3001

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 530 base pairs

TYPE: nucleic acid

```

;
; TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-10-073-353-4

Query Match          9.3%; Score 45.6; DB 9; Length 530;
Best Local Similarity 54.3%; Pred. No. 0.0049;
Matches 113; Conservative 0; Mismatches 94; Indels 1; Gaps 1

QY      22 TCGGGCCGCAATCTTCGCGGGAGATATCTGTCTCTCCTCTCACACGGGCACCCCATAGTGA 81
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      264 TEGACCACATCTCTGCCGCTGCTCGCGCTCATGAGCCGGCCGGGCGGGGGCGG 323
QY      82 CGGCGACGCTCGGGCCCTCTACTGTGGAAAGCTATACATGCTG -ATCGGGCGGCGTGGCGCG 140
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      324 CGGCGGGGGGGGGGGGGGGGGCCCCCAGACAGCGGCCGGGGGGGGCGGTGGCCGCCCGGGCGGGCG 383
QY      141 GGCGGCGGGGTTCTTCGCGCTGTGCGCGGGTGCCGGGTCCGTGGCGGCTTGGCGTCCGCGGCG 200
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      384 GGCGGCGCGGGCGCGCGCGGGCCCCCGGGGGGAGGCGGCGGTGGCGGCCCGCGGGGGGCGGCGCG 443
QY      201 GGCGGCGGATGAGGGCGGCGACCTGGGGTGG 228
        ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB      444 GGCGGCGGGGGCGGGCGGCTCTCTGGGCGG 471

```

```

US-09-825-288A-1/c
: Sequence 1, Application US/09825288A
: Publication No. US20020192822A1
: GENERAL INFORMATION:
: APPLICANT: LEOPARDI, ROSARIO
: APPLICANT: ROIZMAN, BERNARD
: TITLE OF INVENTION: HERPES SIMPLEX VIRUS ICP4 IS AN INHIBITOR OF APOPTOSIS
: FILE REFERENCE: ARD:317USC1
: CURRENT APPLICATION NUMBER: US/09/825,288A
: PRIOR FILING DATE: 2001-04-02
: PRIOR APPLICATION NUMBER: 09/253,821
: PRIOR FILING DATE: 1999-03-01
: PRIOR APPLICATION NUMBER: 08/690,473
: PRIOR FILING DATE: 1996-07-26
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1
: LENGTH: 4257
: TYPE: DNA
: ORGANISM: HERPES VIRUS, TYPE 1
US-09-825-288A-1

Query Match      8.9%  Score 44;  DB 9;  Length 4257;
Match Local Similarity 50.5%;  Pred. No. 0.021;
Matches 107;  Conservative 0;  Mismatches 105;  Indels 0;  Gaps 0.

QY      19  CCATCGGCCGCAACTCTCGGCGGATATCTCTCTCTCTCTCAACGGACCCCATGG 78
Db      613  CCAGCTGTGCGGGCGAGACGGCGTCCCGCGGTCTCTGCGCGCGTGGTGGCCGCGG 554

QY      79  TAGCGGCACAGCTCGCGCCCTCGTGGGAAAGCTGTACATGCTGATCGCGCGGTGCGTCC 138
Db      553  GGGGCCCTCCGTCGCCGCGGGGCGTGTGAAGTCTGTGGGGGTGGTGGGGTGTGGTCGG 494

QY      139  CGGCGCGCGGGCTCTCGCCTCTGTCGGCGGTCGCGGTGCGTGGCGCTTGGCGTCGCGG 198
Db      493  GGTGCTCCCGCCCTCTCTCGTCTCGTCCGCGCCCGCCACCCGAGGGCCCCCGCTGTGGCGT 434

QY      199  CGGCGCGGATGAGGGCGGACCTGTGGTGGTGG 230
Db      433  CTGGGCTCGGGGTGGGCGGCGCGCCGCTGGTGG 402

```

```

Publication No. US20020192778A1
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James
APPLICANT: Molnar, Istvan
APPLICANT: Zickler, Ross
APPLICANT: Cyr, Devon
APPLICANT: Goerlach, Joern
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/10/014.717
CURRENT FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US/09/335.409
PRIOR FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 68750
TYPE: DNA
ORGANISM: Sorangium cellulosum
US-10-014-717-1

Query Match      8.9%;   Score 43.8;   DB 9;   Length 68750;
Best Local Similarity 52.5%;   Pred. No. 0.046;
Matches 96; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

```

```

Y      14  CTTCACCAATGGCGGCAGACTCTGTGSGGGATATTCACGTCGTCCTCCTCCACCGCACCCC   73
Db    63905 CTCACCGAGCGGCAAGTTCTGTCTGTCTGCTCTGCTCGTGGCCCTTGCGCTGTCTCTGTACCGGG   63964
QY     74  CATGTACACGGGCACACTCGGCCCTGCTGTGGGAAGCTGTATCATGCTGATTCGGCGCGCTC   133
Db    63965 CGCGCTTCGTGGCGAGCTCGCGCGGCGGCTGCGCCAGCCGAGGTGCTGCGAGCTCTTC   64024
QY     134 GGTGCGGCGCGCGCGGTCTTCGCGCCTGCTGTGGCGGGTGC CGGTTCGTCGGCGCCTTGGCGCTC   193
Db    64025 GGCGGCGTGGTGTCTGTGGCGCGCTGCTGTGCGGCGCGCTGCTCTGTGGTTCATCGAGTTC   64084
QY     194 CGC   196
Db    64085 CTC   64087

RESULT 4
US-09-712-363-93
; Sequence 93, Application US/09712363
; Patent No. US20020164588A1
GENERAL INFORMATION:
APPLICANT: Eisenberg, David
APPLICANT: Rotstein, Sergio H.
APPLICANT: Marcotte, Edward M.
TITLE OF INVENTION: DETERMINING THE FUNCTIONS AND
FILE REFERENCE: 07419-032001
CURRENT FILING DATE: 2000-11-13
PRIOR APPLICATION NUMBER: US/09/712,363
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: PCT/US00/02246
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: 60/179,531
PRIOR FILING DATE: 2000-02-01
PRIOR APPLICATION NUMBER: 60/117,844
PRIOR FILING DATE: 1999-01-29
PRIOR APPLICATION NUMBER: 60/118,206,
PRIOR FILING DATE: 1999-02-01
PRIOR APPLICATION NUMBER: 60/126,593
PRIOR FILING DATE: 1999-03-26
PRIOR APPLICATION NUMBER: 60/134,093
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 60/134,092
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 60/165,124
PRIOR FILING DATE: 1999-11-12
PRIOR APPLICATION NUMBER: 60/165,086

```



```

1  PRIOR APPLICATION NUMBER: 60/191,078
2  PRIOR FILING DATE: 2000-03-21
3  PRIOR APPLICATION NUMBER: 60/206,848
4  PRIOR FILING DATE: 2000-05-23
5  PRIOR APPLICATION NUMBER: 60/207,727
6  PRIOR FILING DATE: 2000-05-26
7  PRIOR APPLICATION NUMBER: 60/242,578
8  PRIOR FILING DATE: 2000-10-23
9  PRIOR APPLICATION NUMBER: 60/253,625
10 PRIOR FILING DATE: 2000-11-27
11 PRIOR APPLICATION NUMBER: 60/257,931
12 PRIOR FILING DATE: 2000-12-22
13 PRIOR APPLICATION NUMBER: 60/269,308
14 PRIOR FILING DATE: 2001-02-16
15 NUMBER OF SEQ ID NOS: 14110
16 SOFTWARE: FASTSEQ for Windows Version 4.0
17 SEQ ID NO 7935
18 LENGTH: 1665
19 TYPE: DNA
20 ORGANISM: Pseudomonas aeruginosa
21 FEATURE:
22 NAME/KEY: CDS
23 LOCATION: (1)..(1665)
24 OS-09-815-242-7935

```

```

1 RESULT 8
2 US-09-921-232-5/C
3 : Sequence 5, Application US/09921232
4 : Patent No. US20020102681A1
5 :
6 : GENERAL INFORMATION:
7 :
8 : APPLICANT: Martino-Catt, Susan J.
9 : APPLICANT: Wang, Hongyu
10 : APPLICANT: Beach, Larry R.
11 : TITLE OF INVENTION: Polypeptides Controlling Phytate
12 : TITLE OF INVENTION: Metabolism in Plants
13 : FILE REFERENCE: 0706d4
14 :
15 : CURRENT APPLICATION NUMBER: US/09/921,232
16 : CURRENT FILING DATE: 2001-08-02
17 : PRIOR APPLICATION NUMBER: 60/055,446
18 : PRIOR FILING DATE: 1997-08-11
19 : PRIOR APPLICATION NUMBER: 60/055,526
20 : PRIOR FILING DATE: 1997-08-08
21 : PRIOR APPLICATION NUMBER: 60/053,944
22 : PRIOR FILING DATE: 1997-07-28
23 : PRIOR APPLICATION NUMBER: 09/118,442
24 : PRIOR FILING DATE: 1998-07-17
25 : PRIOR APPLICATION NUMBER: 09/677,064
26 : PRIOR FILING DATE: 2000-09-29
27 :
28 : NUMBER OF SEQ ID NOS: 31
29 :
30 : SOFTWARE: FastSeq for Windows Version 3.0
31 : SEQ ID NO 5
32 :
33 : LENGTH: 1428
34 : TYPE: DNA
35 :
36 : ORGANISM: zea mays
37 :
38 : FEATURE:
39 :
40 : NAME/KEY: CDS

```

```

; LOCATION: (118) ... (1176)
US-09-921-232-5

Query Match      8.2%; Score 40.2; DB 10; Length 1428;
Best Local Similarity 49.3%; Pred. No. 0.13;
Matches 105; Conservative 0; Mismatches 108; Indels 0; Gaps 0

Oy 13 TCTTACACATCGCGCGGCACTCTCTCGGAGATATCTCTGTCGTCCTTCATCCAGCGGACCC 72
Db 298 TCTCTGGCGAGCCCCCGGAGCTTGGCTGGAGGAAGCTCTTGGCTTCTGTCGTGAGCG 239

Oy 73 CCATGATAGCGGCGACGCTCGCGCCCTGCTCGGAGAAAGCTATCATGCTGATCGCGCGCT 132
Db 238 CGTACCCACACACGACACTCTCCCGCGGAGGAGCGCGGACACACAGCGCTCTCTCGCGCGCT 179

Oy 133 CGGTGCGCGGCGCGGCGGCTTCTCCGCTGCTCGGCGGTGCGGCTCGGTGGCGCT 192
Db 178 CCGGAGGCGCGCCCGCGCGGCTCGGCTCTCCCTCTCGTCTCCACGACCGCGCAGACCA 119

Oy 193 CCGGCGGCGCGCGATGAGGCGGCGACCTGGG 225
Db 118 TCGCGGCGGACCGACGAGAGCGCACTAGCGGCG 86

```

```

RESULT 9
US-09-921-330-5/c
: Sequence 5, Application US/09921330
: Patent No. US20020102682A1
: GENERAL INFORMATION:
: APPLICANT: Martino-Calt, Susan J.
: APPLICANT: Wang, Hongyu
: APPLICANT: Beach, Larry R.
: TITLE OF INVENTION: Polypeptides Controlling Phytate
: TITLE OF INVENTION: Metabolism in Plants
: FILE REFERENCE: 070603
: CURRENT APPLICATION NUMBER: US/09/921,330
: CURRENT FILING DATE: 2001-08-02
: PRIOR APPLICATION NUMBER: 60/055,446
: PRIOR FILING DATE: 1997-08-11
: PRIOR APPLICATION NUMBER: 60/055,526
: PRIOR FILING DATE: 1997-08-08
: PRIOR APPLICATION NUMBER: 60/053,944
: PRIOR FILING DATE: 1997-07-28
: PRIOR APPLICATION NUMBER: 09/118,442
: PRIOR FILING DATE: 1998-07-17
: PRIOR APPLICATION NUMBER: 09/677,064
: PRIOR FILING DATE: 2000-09-29
: NUMBER OF SEQ ID NOS: 31
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 5
: LENGTH: 1428
: TYPE: DNA
: ORGANISM: Zea mays
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (118)...(1176)
US-09-921-330-5

Query Match      8.2%; Score 40.2; DB 10; Length 1428;
Best Local Similarity 49.3%; Pred. No. 0.15;
Matches 105; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY      13 TCTTCACATAGCGCGGCACAACTCCTCGGGGATTCCTCGTGCCTCTCTCCACACGGCACC 72
Db      298 TCTTGGCCAGGCCCCCGGAGCTTGGCTGGAGAACTCTTGCCCTTCTTCGTGTGAGCG 239

QY      73 CCATGCTAGCGGCCACGCTCGCCCTTGCCCTGGGAAGCTGTACATGCTGATGGCGGCGT 132
Db      238 GGTACCCACACACGACACCTCCGCGGGGAGGCGCGGCACACACGGGCTCTCCGCGGCT 179

QY      133 CGGTCCGCGCGCGCGGCTCTTCGCTGTGTCGGCGGTGCGGCTCCGTGCGGCGCTTGGCGT 192
Db      178 CCGGAGGCGCGCGCGCGGCTCTTCGCTGTGTCGGCGGTGCGGCTCCGCTCCACACGACCGCGCAGACACCA 119

```



```

; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/517,849
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: US 08/979,608
; PRIOR FILING DATE: 1997-11-26
; PRIOR APPLICATION NUMBER: US 60/031,930
; PRIOR FILING DATE: 1996-11-27
; PRIOR APPLICATION NUMBER: US 60/048,547
; PRIOR FILING DATE: 1997-06-03
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48
; LENGTH: 2561
; TYPE: DNA
; ORGANISM: Oryctolagus cuniculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (246)...(1895)
US-10-023-529-48

```

```

Query Match      8.0%; Score 39.4; DB 12; Length 2561;
Best Local Similarity 57.2%; Pred. No. 0.28;
Matches 91; Conservative 0; Mismatches 66; Indels 2; Gaps 1;

```

```

QY 49 GGTCTCTCTCTCCACCGGACACCCCATGTAGCGGCCAGCTCGGCGCCCTGCTGGGAAA 108
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 1032 CTTCCTCCGAGGCTACGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 973
QY 109 GCTGTACATGCTGATCGGCGGCGGTCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCG 168
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 972 GCGGCTGTGGCTG--CTGCGGCGGCGGCGGCGGCGGCGGCGGCTGCTGCTGTGTGGCGG 915
QY 169 TGCCGATCCGTCGCGGCTTGGCGCTCCGCGGCGCGCGCGCGCGCGCGCGCGCGCG 207
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 914 CGCCGCGGCGCTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 876

```

Search completed: December 26, 2002, 08:20:43
 Job time : 316 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2002, 03:47:00 ; Search time 264 Seconds
(without alignments)
110.894 Million cell updates/sec

Title: US-09-548-449-8
Perfect score: 13
Sequence: 1 aatcaannantta 13

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues
Total number of hits satisfying chosen parameters: 2390332

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10	76.9	13	AAA95512	Consensus IHF bind
2	10	76.9	13	ABA02585	IHF binding site c
3	10	76.9	13	ABC22980	Oligonucleotide SE
4	10	76.9	13	ABC22981	Oligonucleotide SE
5	10	76.9	13	ABC22982	Oligonucleotide SE
6	10	76.9	13	ABC22983	Oligonucleotide SE
7	10	76.9	13	ABF44524	Oligonucleotide SE
8	10	76.9	13	ABF44525	Oligonucleotide SE
9	10	76.9	13	ABF44526	Oligonucleotide SE

10	10	76.9	13	23	ABF44527	Oligonucleotide SE
C 11	10	76.9	13	23	ABH60452	Oligonucleotide SE
12	10	76.9	13	23	ABH60453	Oligonucleotide SE
C 13	10	76.9	13	23	ABH60454	Oligonucleotide SE
C 14	10	76.9	13	23	ABH60455	Oligonucleotide SE
C 15	10	76.9	18	20	AAA22841	Integrin subunit b
16	10	76.9	19	21	AAA85862	Cyclin B1 ribozyme
17	10	76.9	19	22	AAH61024	Forward PCR primer
18	10	76.9	24	22	AAH31589	Real-time validati
19	10	76.9	26	24	ABN89968	E. coli K123 tpia
C 20	10	76.9	28	22	AAH82012	tpia gene PCR prim
C 21	10	76.9	28	22	AAH28565	Fungal PCR primer S
C 22	10	76.9	29	24	ABN99540	Paecilomyces vario
C 23	10	76.9	29	24	ABK29137	Tobacco PAPR probe
C 24	10	76.9	36	19	AAV09780	Tobacco PAPR probe
C 25	10	76.9	36	19	AAV09780	Tobacco PAPR probe
C 26	10	76.9	43	19	AAV09787	Synthetic plasmid
C 27	10	76.9	50	20	AAH52183	Synthetic plasmid
C 28	10	76.9	50	20	AAH52050	Human DNA containi
C 29	10	76.9	50	22	AAH79643	Human DNA containi
C 30	10	76.9	51	22	AAH79644	Human SNP oligonuc
C 31	10	76.9	51	22	AAH32103	Leucine zipper-hum
C 32	10	76.9	53	24	ABA92528	Human ovarian tran
C 33	10	76.9	60	24	ABN41535	Human gene signatu
C 34	10	76.9	62	22	AAH24562	Tetracycline promo
C 35	10	76.9	78	16	AAH19807	HBV infection regu
C 36	10	76.9	87	22	AAH23585	Staphylococcus aur
C 37	10	76.9	87	24	AAH34618	3'-5' m1cF RNA str
C 38	10	76.9	89	18	AAV77658	Human secreted pro
C 39	10	76.9	96	14	AAQ41243	Candida albicans-s
C 40	10	76.9	97	21	AAH17769	Oligonucleotide SE
C 41	10	76.9	100	15	AAH62573	Oligonucleotide SE
C 42	9.6	73.8	13	23	ABF19650	Oligonucleotide SE
C 43	9.6	73.8	13	23	ABF19651	Oligonucleotide SE
C 44	9.6	73.8	13	23	ABF19652	Oligonucleotide SE
C 45	9.6	73.8	13	23	ABF19653	Oligonucleotide SE

ALIGNMENTS

RESULT 1	
AAA95512	AAA95512 standard; DNA; 13 BP.
AC	AAA95512;
XX	
XX	27-FEB-2001 (first entry)
DT	
XX	
DE	Consensus IHF binding site.
XX	
KW	Bacteria; infection; drug-resistant pathogen; cancer; typhoid fever;
KW	bacterial meningitis; tuberculosis; antisense strand; ribozyme;
KW	toxic protein; ds.
XX	
OS	Unidentified.
XX	
PN	W0200061804-A1.
XX	
PD	19-OCT-2000.
XX	
PF	14-APR-2000; 2000WO-US10229.
XX	
PR	14-APR-1999; 99US-0291902.
XX	
PA	13-APR-2000; 2000US-0548449.
XX	
PA	(MUSC-) MUSC FOUND RES DEV.
XX	(DYPE-) UNIV PENN STATE.
PI	Norris J, Clawson G, Westwater C, Schofield D, Schmidt M, Hoel B;
XX	Dolan J, Pan W;
DR	WPI; 2000-638570/61.

XX New recombinant nucleic acid comprising a nucleotide sequence encoding
PT one or more toxic agents operably linked to a pathogen specific or
PT tissue specific promoter is useful for inhibiting replication of a
PT pathogen -
XX
XX Example; Page 67; 113pp; English.
XX
XX The present sequence comprises the consensus IHF binding site, which
CC was used in the construction of the pathogen and tissue-specific toxic
CC agents of the invention. These agents may be antisense sequences,
CC ribozymes or toxic proteins. These agents can be used in the treatment of cancer
CC and infections including hepatitis, herpes, malaria, bacterial
CC meningitis, typhoid fever and tuberculosis.
XX
XX Sequence 13 BP; 6 A; 1 C; 0 G; 3 T; 3 other;
S0

Query Match 76.9%; Score 10; DB 21; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.1e+04;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

OY 1 AATCAANNANTTA 13
|||||
DB 1 AATCAANNANTTA 13

RESULT 2
ABA02585-standAid.DNA-13-BP.
XX ABA02585;
AC
XX
DT 05-FEB-2002 (first entry)
XX
DE IHF binding site consensus sequence.
XX
KW Infection; antisense RNA; ribozyme; DNzyme; antiviral; gene therapy;
KW papilloma virus; hepatitis B virus; cytotoxic; cytosstatic; wart;
KW cervical dysplasia; cervical carcinoma; carcinoma; laryngeal papilloma;
KW IHF binding site; ss.
XX
OS Synthetic.
XX
W0200179524-A2.
XX
PN 25-OCT-2001.
BD
XX
EF 13-APR-2001; 2001WO-US12130.
XX
PR 13-APR-2000; 2000US-0548449.
PR 07-DEC-2000; 2000US-251810P.
XX
PA (UYSC-) UNIV SOUTH CAROLINA.
PA (PENNC-) PENN STATE RES FOUND.
XX
XX Norris JS, Clawson GA, Westwater C, Schofield D, Schmidt MG;
PI Hoel B, Dolan J, Pan W;
PI WPI: 2001-607700/69.
DR

Novel nucleic acid for the treatment of papilloma or hepatitis virus
induced conditions comprises a catalytic region which produces a
cytotoxic or cytostatic effect in the infected cell -
Examples; Page 79; 143pp; English.

The invention relates to the discovery, identification and
characterisation of toxic agents lethal to pathogens and methods for
targeting such toxic agents to a pathogen or pathogen infected cells in
order to treat and/or eradicate the infection. In particular the
invention relates to at least one nucleic acid molecule, which
specifically hybridises to mRNA encoding at least one vital protein
associated with the transformation or plasmid copy number control, which

CC	hybridises to a viral polyadenylation signal or a core, pre-core or
CC	polymerase encoding sequence. Specifically, the invention relates to the
CC	delivery of one or more toxic gene products, antisense RNAs, ribozymes,
CC	DNAzymes or a combination thereof. The nucleic acids have antiviral
CC	activity and can be used in gene therapy. They are useful for the
CC	treatment of papilloma or hepatitis virus induced conditions and can
CC	produce a cytotoxic or cytostatic effect in papillomavirus or hepatitis B
CC	infected cells. The papilloma virus induced condition is selected from
CC	warts, cervical dysplasia, cervical carcinoma, carcinoma in situ and
CC	laryngeal papilloma. The present sequence is that of a the IMF binding
CC	sile consensus sequence featured in the bacteriophage P1 pac site
CC	sequence (ABA02577).
XX	
SQ	Sequence 13 BP; 6 A; 1 C; 0 G; 3 T; 3 other;
XX	
Query Match	76.9%; Score 10; DB 22; Length 13;
Best Local Similarity	100.0%; Pred. No. 1,1e+04;
Matches 13; Conservative	0; Mismatches 0; Indels 0; Gaps 0
OY	1 AATCAANNANTTA 13 1 AATCAANNANTTA 13
Db	
RESULT_3	
ABC22980/c	
ID	ABC22980 standard; DNA; 13 BP.
XX	
AC	ABC22980;
DT	20-FEB-2002 (first entry)
DE	oligonucleotide SEQ ID NO 22997 for detecting SNP TSC0004518.
XX	
KW	SNP: single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KM	peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
XX	central nervous system; gastrointestinal; respiratory; immune; metabolic.
OS	Homo sapiens.
XX	
WO200177384-A2.	
PN	
PD	18-OCT-2001.
PF	06-APR-2001; 2001WO-IB00713.
PR	
PA	07-APR-2000; 2000DE-1019173.
PI	(EPig-) EPIDENOMICS AG.
PB	
PI	Olek A, Piepenbrock C, Berlin K;
DR	WPI; 2001-657177/75.
PT	
PT	Set of oligonucleotides, useful for diagnosis and cell typing, is
PT	designed to detect single nucleotide polymorphisms and cytosine
PT	methylation status -
PS	
PS	Claim 1; SEQ ID 22997; 29pp + Sequence Listing; German.
XX	
XX	This invention describes novel oligonucleotide primers or peptide nucleic
CC	acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC	and cytosine methylation status in chemically pretreated genomic DNA. The
CC	oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC	range of diseases including immune system, gastrointestinal, respiratory,
CC	central nervous system, cardiovascular and metabolic disorders. The
CC	oligomers are also used for detecting cell type differentiation.
CC	ABG00010-ABG9989, ABF00010-ABF9989, ABH00010-ABH9989 and
CC	ABI00010-ABI82073 represent the oligomers described in the invention.
CC	NOTE: The sequence data for this patent did not form part of the printed
CC	specification, but was obtained in electronic format from WIPO at
CC	ftp.wipo.int/pub/published_pct_sequences.
XX	

SO Sequence 13 BP; 4 A; 0 C; 2 G; 7 T; 0 other;
Query Match 76.9%; Score 10; DB 23; Length 13;
Best Local Similarity 76.9%; Pred. No. 1.1e+04;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 AATCAANNANTTA 13
13 AATCAATAACTTA 1

Db 13 AATCAATAACTTA 1

RESULT 4
ABC22981
ID ABC22981 standard; DNA; 13 BP.
AC ABC22981;
XX
XX 20-FEB-2002 (first entry)
DT
XX
DE Oligonucleotide SEQ ID NO 22998 for detecting SNP TSC0004518.
XX
XX SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
XX Homo sapiens.
XX
XX WO200177384-A2.
XX
XX 18-OCT-2001.
XX
XX 06-APR-2001; 2001WO-IB00713.
XX
XX 07-APR-2000; 2000DE-1019173.
XX
XX (EPIC-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX
XX WPI; 2001-657177/75.
XX
XX Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status
PS
PS Claim 1; SEQ ID 22998; 29pp + Sequence Listing: German.
XX
XX This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC AB100010-AB182073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX
XX Sequence 13 BP; 7 A; 2 C; 0 G; 4 T; 0 other;
SO

Query Match 76.9%; Score 10; DB 23; Length 13;
Best Local Similarity 76.9%; Pred. No. 1.1e+04;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 AATCAANNANTTA 13
13 AATCAATAACTTA 1

Db 1 AATCAATAACTTA 13

RESULT 5
ABC22982/c

ID ABC22982 standard; DNA; 13 BP.
XX
XX ABC22982;
AC
XX 20-FEB-2002 (first entry)
DT
XX
XX Oligonucleotide SEQ ID NO 22999 for detecting SNP TSC0004518.
DE
XX
XX SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
XX Homo sapiens.
XX
XX WO200177384-A2.
XX
XX 18-OCT-2001.
XX
XX 06-APR-2001; 2001WO-IB00713.
XX
XX 07-APR-2000; 2000DE-1019173.
XX
XX (EPIC-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX
XX WPI; 2001-657177/75.
XX
XX Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status
PS
PS Claim 1; SEQ ID 22999; 29pp + Sequence Listing: German.
XX
XX This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC AB100010-AB182073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX
XX Sequence 13 BP; 3 A; 0 C; 2 G; 8 T; 0 other;
SO

Query Match 76.9%; Score 10; DB 23; Length 13;
Best Local Similarity 76.9%; Pred. No. 1.1e+04;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 AATCAANNANTTA 13
13 AATCAATAACTTA 1

Db 13 AATCAATAACTTA 1

RESULT 6
ABC22983
ID ABC22983 standard; DNA; 13 BP.
AC ABC22983;
XX
XX 20-FEB-2002 (first entry)
DT
XX
XX Oligonucleotide SEQ ID NO 23000 for detecting SNP TSC0004518.
DE
XX
XX SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
XX Homo sapiens.
OS

```
XX WO200177384-A2.
XX
XX 18-OCT-2001.
XX
XX 06-APR-2001; 2001WO-IB00713.
XX
XX 07-APR-2000; 2000DE-1019173.
XX
XX (EPiG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX
XX WPI; 2001-657177/75.
XX
XX Set of oligonucleotides, useful for diagnosis and cell typing, is
XX designed to detect single nucleotide polymorphisms and cytosine
XX methylation status -
XX
XX Claim 1; SEQ ID 23000; 29pp + Sequence Listing; German.
XX
XX This invention describes novel oligonucleotide primers or peptide nucleic
XX acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
XX and cytosine methylation status in chemically pretreated genomic DNA. The
XX oligonucleotides are used for diagnosis and/or prognosis of cancer and a
XX range of diseases including immune system and metabolic disorders. The
XX central nervous system, cardiovascular and gastrointestinal, respiratory,
XX oligomers are also used for detecting cell type differentiation.
XX ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
XX ABI00010-ABI82073 represent the oligomers described in the invention.
XX NOTE: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 13 BP; 8 A; 2 C; 0 G; 3 T; 0 other;
XX
SQ
XX
XX Query Match 76.9%; Score 10; DB 23; Length 13;
XX Best Local Similarity 76.9%; Pred. No. 1.1e+04;
XX Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
XX 1 AATCAANNANTTA 13
XX ||||| |||
XX 1 AATCAAAACTTA 13
XX
DB
XX
XX RESULT 7
XX ABF44524/C
XX ID ABF44524 standard; DNA; 13 BP.
XX
XX ABR44524;
XX
XX 21-FEB-2002 (first entry)
XX
XX Oligonucleotide SEQ ID NO 144521 for detecting SNP TSC0036340.
XX
XX SNP: single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
XX peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
XX central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX Homo sapiens.
XX
XX WO200177384-A2.
XX
XX 18-OCT-2001.
XX
XX 06-APR-2001; 2001WO-IB00713.
XX
XX 07-APR-2000; 2000DE-1019173.
XX
XX (EPiG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX
XX
```

```
DR WPI; 2001-657177/75.
XX
XX Set of oligonucleotides, useful for diagnosis and cell typing, is
XX designed to detect single nucleotide polymorphisms and cytosine
XX methylation status -
XX
XX Claim 1; SEQ ID 144521; 29pp + Sequence Listing; German.
XX
XX This invention describes novel oligonucleotide primers or peptide nucleic
XX acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
XX and cytosine methylation status in chemically pretreated genomic DNA. The
XX oligonucleotides are used for diagnosis and/or prognosis of cancer and a
XX range of diseases including immune system, gastrointestinal, respiratory,
XX central nervous system, cardiovascular and metabolic disorders. The
XX oligomers are also used for detecting cell type differentiation.
XX ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
XX ABI00010-ABI82073 represent the oligomers described in the invention.
XX NOTE: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 13 BP; 4 A; 0 C; 1 G; 8 T; 0 other;
XX
SQ
XX
XX Query Match 76.9%; Score 10; DB 23; Length 13;
XX Best Local Similarity 76.9%; Pred. No. 1.1e+04;
XX Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
XX 1 AATCAANNANTTA 13
XX ||||| |||
XX 13 AATCAATAATTA 1
XX
DB
XX
XX RESULT 8
XX ABR44525
XX ID ABR44525 standard; DNA; 13 BP.
XX
XX ABR44525;
XX
XX 21-FEB-2002 (first entry)
XX
XX Oligonucleotide SEQ ID NO 144522 for detecting SNP TSC0036340.
XX
XX SNP: single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
XX peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
XX central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX Homo sapiens.
XX
XX WO200177384-A2.
XX
XX 18-OCT-2001.
XX
XX 06-APR-2001; 2001WO-IB00713.
XX
XX 07-APR-2000; 2000DE-1019173.
XX
XX (EPiG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX
XX WPI; 2001-657177/75.
XX
XX Set of oligonucleotides, useful for diagnosis and cell typing, is
XX designed to detect single nucleotide polymorphisms and cytosine
XX methylation status -
XX
XX Claim 1; SEQ ID 144522; 29pp + Sequence Listing; German.
XX
XX This invention describes novel oligonucleotide primers or peptide nucleic
XX acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
XX and cytosine methylation status in chemically pretreated genomic DNA. The
XX oligonucleotides are used for diagnosis and/or prognosis of cancer and a
XX range of diseases including immune system, gastrointestinal, respiratory,
```

CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989 and
CC AB100010-AB182073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 13 BP; 8 A; 1 C; 0 G; 4 T; 0 other;
Query Match 76.9%; Score 10; DB 23; Length 13;
Best Local Similarity 76.9%; Pred. No. 1.1e+04;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AATCAANNANTTA 13
Db 1 AATCAATTAATTA 13
|||||
RESULT 9
ABF44526/c
ID ABF44526 standard; DNA; 13 BP.
XX
AC ABF44526;
XX
DT 21-FEB-2002 (first entry)
XX
DE Oligonucleotide SEQ ID NO 144523 for detecting SNP TSC0036340.
XX
KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS Homo sapiens.
XX
PN W0200177384-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-1B00713.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PS Claim 1; SEQ ID 144523; 29pp + Sequence listing; German.
XX
SQ Sequence 13 BP; 3 A; 0 C; 2 G; 8 T; 0 other;
Query Match 76.9%; Score 10; DB 23; Length 13;
Best Local Similarity 76.9%; Pred. No. 1.1e+04;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AATCAANNANTTA 13
Db 1 AATCAATTAATTA 13
|||||
RESULT 10
ABF44527
ID ABF44527 standard; DNA; 13 BP.
XX
AC ABF44527;
XX
DT 21-FEB-2002 (first entry)
XX
DE Oligonucleotide SEQ ID NO 144524 for detecting SNP TSC0036340.
XX
KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS Homo sapiens.
XX
PN W0200177384-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-1B00713.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PS Claim 1; SEQ ID 144524; 29pp + Sequence listing; German.
XX
SQ Sequence 13 BP; 8 A; 2 C; 0 G; 3 T; 0 other;
Query Match 76.9%; Score 10; DB 23; Length 13;
Best Local Similarity 76.9%; Pred. No. 1.1e+04;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AATCAANNANTTA 13
Db 1 AATCAATTAATTA 13
|||||
RESULT 11
ABH60452/c
ID ABH60452 standard; DNA; 13 BP.
XX
AC ABH60452;
XX
DT 22-FEB-2002 (first entry)

QY 1 AATCAANNANTTA 13
Db 13 AATCAACAAATTA 1
|||||
RESULT 10
ABF44527
ID ABF44527 standard; DNA; 13 BP.
XX
AC ABF44527;
XX
DT 21-FEB-2002 (first entry)
XX
DE Oligonucleotide SEQ ID NO 144524 for detecting SNP TSC0036340.
XX
KW SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS Homo sapiens.
XX
PN W0200177384-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-1B00713.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PS Claim 1; SEQ ID 144524; 29pp + Sequence listing; German.
XX
SQ Sequence 13 BP; 8 A; 2 C; 0 G; 3 T; 0 other;
Query Match 76.9%; Score 10; DB 23; Length 13;
Best Local Similarity 76.9%; Pred. No. 1.1e+04;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AATCAANNANTTA 13
Db 1 AATCAACAAATTA 13
|||||
RESULT 11
ABH60452/c
ID ABH60452 standard; DNA; 13 BP.
XX
AC ABH60452;
XX
DT 22-FEB-2002 (first entry)

XX Oligonucleotide SEQ ID NO 260429 for detecting SNP TSC0063245.
DE
XX SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS Homo sapiens.
XX
PN WO200177384-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-1B00713.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PA (EPIC-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-657177/75.
XX
PT Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status -
XX
PS Claim 1; SEQ ID 260429; 29pp + Sequence Listing; German.
XX
CC This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC AB100010-AB182073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 13 BP; 5 A; 0 C; 1 G; 7 T; 0 other;
XX
Query Match 76.9%; Score 10; DB 23; Length 13;
Best Local Similarity 76.9%; Pred. No. 1.1e+04;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
OY 1 AATCAANNANTTA 13
DB 13 AATCAANNANTTTTA 1
XX
RESULT 12
ABH60453
ID ABH60453 standard; DNA; 13 BP.
XX
AC ABH60453;
XX
XX 22-FEB-2002 (first entry)
XX
DE Oligonucleotide SEQ ID NO 260430 for detecting SNP TSC0063245.
XX
XX SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS Homo sapiens.
XX
PN WO200177384-A2.
XX
PD 18-OCT-2001.
XX
PT

PF 06-APR-2001; 2001WO-1B00713.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PA (EPIC-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-657177/75.
XX
PT Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status -
XX
PS Claim 1; SEQ ID 260430; 29pp + Sequence Listing; German.
XX
CC This invention describes novel oligonucleotide primers or peptide nucleic
CC acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP)
CC and cytosine methylation status in chemically pretreated genomic DNA. The
CC oligonucleotides are used for diagnosis and/or prognosis of cancer and a
CC range of diseases including immune system, gastrointestinal, respiratory,
CC central nervous system, cardiovascular and metabolic disorders. The
CC oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC AB100010-AB182073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 13 BP; 7 A; 1 C; 0 G; 5 T; 0 other;
XX
Query Match 76.9%; Score 10; DB 23; Length 13;
Best Local Similarity 76.9%; Pred. No. 1.1e+04;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
OY 1 AATCAANNANTTA 13
DB 1 AATCAANNANTTTTA 13
XX
RESULT 13
ABH60454/C
ID ABH60454 standard; DNA; 13 BP.
XX
AC ABH60454;
XX
XX 22-FEB-2002 (first entry)
XX
DE Oligonucleotide SEQ ID NO 260431 for detecting SNP TSC0063245.
XX
XX SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
OS Homo sapiens.
XX
PN WO200177384-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-1B00713.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PA (EPIC-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-657177/75.
XX
PT Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status -
XX

XX Claim 1; SEQ ID 260431; 29pp + Sequence Listing; German.
XX
XX This invention describes novel oligonucleotide primers or peptide nucleic acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP) and cytosine methylation status in chemically pretreated genomic DNA. The oligonucleotides are used for diagnosis and/or prognosis of cancer and a range of diseases including immune system, gastrointestinal, respiratory, central nervous system, cardiovascular and metabolic disorders. The oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC AB100010-AB182073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format from WIPO at [ftp.wipo.int/pub/published_pcl_sequences](http://wipo.int/pub/published_pcl_sequences).
CC
CC
SQ Sequence 13 BP; 5 A; 0 C; 2 G; 6 T; 0 other;
Query Match 76.9%; Score 10; DB 23; Length 13;
Best Local Similarity 76.9%; Pred. No. 1.1e+04;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AATCAANNANTTA 13
13 AATCAACTATTTA 1
DB 13 AATCAACTATTTA 1
RESULT 14
ABH60455
ID ABH60455 standard; DNA; 13 BP.
AC ABH60455;
XX 22-FEB-2002 (first entry)
DE Oligonucleotide SEQ ID NO 260432 for detecting SNP TSC0063245.
XX
XX SNP; single nucleotide polymorphism; human; diagnosis; PNA; cancer; CNS;
KW peptide nucleic acid; cytosine methylation; cardiovascular; primer; ss;
KW central nervous system; gastrointestinal; respiratory; immune; metabolic.
XX
XX Homo sapiens.
XX
XX WO200177384-A2.
XX 18-OCT-2001.
XX
XX 06-APR-2001; 2001WO-IB00713.
XX
XX 07-APR-2000; 2000DE-1019173.
XX
XX (EPIC-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX WPI; 2001-657177/75.
XX
XX Set of oligonucleotides, useful for diagnosis and cell typing, is
PT designed to detect single nucleotide polymorphisms and cytosine
PT methylation status
XX
XX Claim 1; SEQ ID 260432; 29pp + Sequence Listing; German.
XX
XX This invention describes novel oligonucleotide primers or peptide nucleic acid (PNA) oligomers for detecting single nucleotide polymorphisms (SNP) and cytosine methylation status in chemically pretreated genomic DNA. The oligonucleotides are used for diagnosis and/or prognosis of cancer and a range of diseases including immune system, gastrointestinal, respiratory, central nervous system, cardiovascular and metabolic disorders. The oligomers are also used for detecting cell type differentiation.
CC ABC00010-ABC99989, ABF00010-ABF99989, ABH00010-ABH99989 and
CC AB100010-AB182073 represent the oligomers described in the invention.
CC NOTE: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format from WIPO at
CC [ftp.wipo.int/pub/published_pcl_sequences](http://wipo.int/pub/published_pcl_sequences).
CC
SQ Sequence 13 BP; 6 A; 2 C; 0 G; 5 T; 0 other;
Query Match 76.9%; Score 10; DB 23; Length 13;
Best Local Similarity 76.9%; Pred. No. 1.1e+04;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AATCAANNANTTA 13
1 AATCAACTATTTA 13
DB 1 AATCAACTATTTA 13
RESULT 15
AAA22841/C
ID AAA22841 standard; RNA; 18 BP.
AC AAA22841;
XX 19-JUN-2000 (first entry)
DE Integrin subunit beta 3 substrate sequence SEQ ID NO:6067.
XX
XX Human; aryl hydrocarbon nuclear transport; ARNT; Tie-2; angiogenesis;
KW integrin alpha 6 subunit; integrin subunit beta 3; hairpin ribozyme;
KW hammerhead ribozyme; angiogenic factor; cytosolic; antidiabetic;
KW ophthalmologic; antiinflammatory; antiarthritic; antipsoriatic; ARMD;
KW dermatologic; RNA cleavage; cancer; diabetic retinopathy; arthritis;
KW age related macular degeneration; inflammation; neovascular glaucoma;
KW myopic degeneration; psoriasis; verruca vulgaris; angiodioma;
KW tuberculous sclerosis; pot-wine stain; Sturge Weber syndrome;
KW Kippel-Trenaunay-Weber syndrome; Osler-Weber-Rendu syndrome; ss.
XX
XX Homo sapiens.
XX
XX WO9950403-A2.
XX
XX 07-OCT-1999.
XX
XX 24-MAR-1999; 99WO-US06507.
XX
XX 27-MAR-1998; 98US-0079678.
XX
XX (RIBO-) RIBOZYME PHARM INC.
XX
XX Payco PA, Roberts E, Jarvis T, Coeshott C, McSwiggen JA;
XX WPI; 1999-591315/50.
XX
XX Novel ribozymes for modulating the synthesis, expression and/or
PT stability of an mRNA encoding an angiogenic factors
XX
XX Claim 54; Page 245; 305pp; English.
XX
XX The present invention describes enzymatic nucleic acid molecules with
CC RNA cleaving activity, which specifically cleave RNA encoded by an aryl
CC hydrocarbon nuclear transporter (ARNT) gene, an integrin subunit beta 3
CC gene, an integrin alpha 6 subunit gene, or a Tie-2 gene. AA16775 to
CC AA17167 and AA17561 to AA17622 represent ribozyme sequences for ARNT,
CC and AA17168 to AA17560 and AA17623 to AA17684 represent their
CC corresponding target sequences; AA17685 to AA18385 and AA19087 to
CC AA19154 represent ribozyme sequences for Tie-2, and AA18386 to AA19086
CC and AA19155 to AA19222 represent their corresponding target sequences;
CC AA19223 to AA20361 and AA21501 to AA21595 represent ribozyme
CC sequences for integrin alpha 6 subunit, and AA20362 to AA21500 and
CC AA21596 to AA21688 represent their corresponding target sequences;
CC AA21689 to AA22475 and AA23263 to AA23342 represent ribozyme sequence
CC for integrin subunit beta 3, and AA22476 to AA23262, AA23343 to
CC AA23422 represent their corresponding target sequences. The ribozymes of
CC the invention are used for modulating the synthesis, expression and/or
CC stability of an mRNA encoding angiogenic factor, especially ARNT,
CC integrin subunit beta 3, integrin subunit alpha 6, or Tie-2. They are

CC especially used to treat cancer, diabetic retinopathy, age related
 CC macular degeneration (AMD), inflammation, and arthritis, as well as
 CC neovascular glaucoma, myopic degeneration, psoriasis, verruca vulgaris,
 CC angiofibroma of tuberous sclerosis, port-wine stains, Sturge Weber
 CC syndrome, Kippel-Trenaunay-Weber syndrome, Osler-Weber-Rendu syndrome,
 CC and other syndromes and diseases related to the levels of ARNT, Tie-2,
 CC integrin subunit alpha-6, or integrin subunit beta-3.

XX Sequence 18 BP; 4 A; 1 C; 1 G; 12 U; 0 other;

Query Match 76.9%; Score 10; DB 20; Length 18;
 Best Local Similarity 76.9%; Pred. No. 1.le+04;

Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 AATCAANNANTTA 13

Db 14 AATCAAAAAATTA 2

Search completed: December 26, 2002, 08:24:21
 Job time : 271 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 26, 2002, 08:15:31 ; Search time 1969 Seconds
(without alignments)
106.928 Million cell updates/sec

Title: US-09-548-449-8
Perfect score: 13
Sequence: 1 aatcaannantta 13

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 segs, 8097743376 residues
Total number of hits satisfying chosen parameters: 357874

Minimum DB seq length: 0
Maximum DB seq length: 100

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:*
1: em_estba:*
2: em_estchum:*
3: em_estlin:*
4: em_estmu:*
5: em_estrov:*
6: em_estropl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pin:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_tod:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10	76.9	21	9	AU258344
2	10	76.9	29	17	BH854703
3	10	76.9	31	17	BH854706
4	10	76.9	34	17	BH854706
5	10	76.9	35	17	AZ439152
6	10	76.9	36	17	BH854728

C 7	10	76.9	37	17	BH854701	BH854701 SALK_0889
C 8	10	76.9	37	17	BH854911	BH854911 SALK_0872
C 9	10	76.9	46	17	AZ359655	AZ359655 1M0102A08
C 10	10	76.9	46	17	AZ508716	AZ508716 1M0351114
C 11	10	76.9	47	17	AZ481968	AZ481968 1M0306110
C 12	10	76.9	49	10	AV856952	AV856952 AV856952
C 13	10	76.9	49	17	BH810378	BH810378 SALK_0492
C 14	10	76.9	49	17	BH861490	BH861490 SALK_0816
C 15	10	76.9	51	9	AU257483	AU257483 AU257483
C 16	10	76.9	51	17	BH791803	BH791803 SALK_0614
C 17	10	76.9	52	9	AU060328	AU060328 AU060328
C 18	10	76.9	55	13	B1946339	B1946339 sr87604.Y
C 19	10	76.9	56	17	AZ363074	AZ363074 1M0108124
C 20	10	76.9	56	17	BH847572	BH847572 SALK_0547
C 21	10	76.9	57	12	BF643255	BF643255 NF004G01E
C 22	10	76.9	57	17	BH847577	BH847577 SALK_0547
C 23	10	76.9	58	14	B0567340	B0567340 g188d04.Y
C 24	10	76.9	59	17	BH854923	BH854923 SALK_0872
C 25	10	76.9	63	9	A1949463	A1949463 wg10d10.X
C 26	10	76.9	65	17	BH813114	BH813114 SALK_0637
C 27	10	76.9	69	13	BM270723	BM270723 sak16q10.
C 28	10	76.9	69	17	BH857895	BH857895 SALK_0875
C 29	10	76.9	70	9	AU013546	AU013546 AU013546
C 30	10	76.9	70	9	AU013844	AU013844 AU013844
C 31	10	76.9	71	17	AZ833202	AZ833202 2M0115E08
C 32	10	76.9	72	13	B1943231	B1943231 sc07a02.Y
C 33	10	76.9	73	9	AA724688	AA724688 a113h09.S
C 34	10	76.9	74	17	AL768789	AL768789 Arabidops
C 35	10	76.9	75	14	N85300	N85300 J3129P Huma
C 36	10	76.9	76	17	AZ853764	AZ853764 2M0157C10
C 37	10	76.9	78	10	AV957827	AV957827 AV957827
C 38	10	76.9	78	14	D19926	D19926 HUMGSO0884
C 39	10	76.9	80	9	AU010897	AU010897 AU010897
C 40	10	76.9	80	17	TA94H050	TA94H050 T. brucei
C 41	10	76.9	81	17	BH853043	BH853043 SALK_0759
C 42	10	76.9	81	17	BH854995	BH854995 SALK_0870
C 43	10	76.9	81	17	BH857864	BH857864 SALK_0875
C 44	10	76.9	82	9	AU010817	AU010817 AU010817
C 45	10	76.9	82	9	AU010818	AU010818 AU010818

ALIGNMENTS

RESULT 1
AU258344
LOCUS
DEFINITION AU258344 3'-directed mouse cDNA library Mus musculus cDNA clone
ACCESSION
VERSION AU258344
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

21 bp MRNA linear EST 25-APR-2002
AU258344 3'-directed mouse cDNA library Mus musculus cDNA clone
BED0012727 3', mRNA sequence.
AU258344
AU258344.1 GI:20323820
EST.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 21)
Kato, K. and Matoba, R.
Generation of expressed sequence tags from mouse brain
Unpublished (2002)
Contact: Kikuya Kato
Graduate School of Biological Sciences
Nara Institute of Science and Technology
8916-5 Takayama, Ikoma, Nara 630-0101, Japan
Tel: 81-743-72-5581
Fax: 81-743-72-5589
Email: kkatob@bs.aist-nara.ac.jp,
URL: http://love2.aist-nara.ac.jp/BBD/index.html.
1. 21
Location/Qualifiers
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="BED0012727"

FEATURES
source

/clone_lib="3'-directed mouse cDNA library"
 /class_type="brain"
 /note="Vector: pGEM-T-easy"
 ORIGIN

Query Match 76.9% Score 10; DB 9; Length 21;
 Best Local Similarity 76.9% Pred. No. 3.5e+04;
 Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AATCAANNANTTA 13
 ||||| | |||
 Db 7 AATCAAAATTTTA 19

RESULT 2
 BH854703/c 29 bp DNA linear GSS 08-JUL-2002
 LOCUS
 DEFINITION Arabidopsis thaliana TDNA insertion lines
 Arabidopsis thaliana genomic clone SALK_088982.46.95.x, DNA

ACCESSION
 VERSION BH854703
 KEYWORDS
 SOURCE

ORGANISM
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 29)
 Alonso,J.M., Leisner,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrihab,
 C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.,
 Zimmerman,J. and Ecker,J.R.
 A Sequence-Indexed Library of Insertion Mutations in the
 Arabidopsis Genome
 Unpublished (2001)
 Contact: Joseph R. Ecker
 Salk Institute Genomic Analysis Laboratory (SIGAL)
 The Salk Institute for Biological Studies
 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
 Tel: 858 453 4100 x1752
 Fax: 858 558 6379
 Email: jecker@salk.edu

TITLE
 JOURNAL
 COMMENT

FEATURES
 source
 This is single pass sequence recovered from the left border of
 TDNA. This sequence lies within 300 bases of the 5' end of
 A5953910.
 Class: TDNA tagged.

Location/Qualifiers
 1..29
 /organism="Arabidopsis thaliana"
 /strain="Columbia 0"
 /db_xref="taxon:3702"
 /clone="SALK_088982.46.95.x"
 /note="PCR was performed on Arabidopsis thaliana lines
 each of which contains one or more TDNA insertion
 elements. The resultant fragment for each line was
 directly sequenced to determine the genomic sequence at
 the site of insertion. Details of the protocols used can
 be found at http://signal.salk.edu/cdna_protocols.html"
 BASE COUNT 11 a 1 c 4 g 13 t
 ORIGIN

Query Match 76.9% Score 10; DB 17; Length 29;
 Best Local Similarity 76.9% Pred. No. 4e+04;
 Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AATCAANNANTTA 13
 ||||| | |||
 Db 15 AATCAAAATTTTA 3

RESULT 3

AZ837501/c 31 bp DNA linear GSS 20-FEB-2001
 LOCUS
 DEFINITION 2M0132M13R Mouse 10kb plasmid UUC1M library Mus musculus genomic
 clone UUCG2M0132M13 R, DNA sequence.

ACCESSION
 VERSION AZ837501
 KEYWORDS
 SOURCE

ORGANISM
 house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 1 (bases 1 to 31)
 Dunn,D., Aoyagi,A., Barber,M., Baacorn,T., Duval,B., Hamll,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Rellly
 M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausen,A.
 and Wright,D., Weiss,R.
 Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 Unpublished (2000)
 Contact: Robert B. Weiss
 University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLG, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0132 row: M column: 13
 Seq primer: CACACGACGAACAGCTATGACC
 Class: plasmid sense stop: 31.
 High quality sequence

FEATURES
 source

Location/Qualifiers
 1..31
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUCG2M0132M13"
 /clone_lib="Mouse 10kb plasmid UUC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (<http://www.jax.org/resources/documents/dnares/>). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pMD42 (g1473211419b1AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the Insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

BASE COUNT 6 a 4 c 9 g 12 t
 ORIGIN

Query Match 76.9% Score 10; DB 17; Length 31;
 Best Local Similarity 76.9% Pred. No. 4.2e+04;
 Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AATCAANNANTTA 13
 ||||| | |||
 Db 27 AATCAAAATTTTA 15

RESULT 4

BH854706/c
LOCUS BH854706 34 bp DNA linear GSS 08-JUL-2002
DEFINITION SALK_088979.46.35.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_088979.46.35.x, DNA
sequence.
ACCESSION BH854706
VERSION BH854706.1 GI:21704296
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadri nab
1 (bases 1 to 34)
AUTHORS 'C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.,
Zimmerman,J., and Ecker,J.R.
A sequence-indexed library of insertion mutations in the
Arabidopsis Genome
Unpublished (2001)
JOURNAL Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGNAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA. This sequence lies within 300 bases of the 5' end of
At5g53910.
Class: TDNA tagged.
FEATURES
source
Location/Qualifiers
1..34
/organism="Arabidopsis thaliana"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_088979.46.35.x"
/note="Arabidopsis thaliana TDNA insertion lines"
/clone_1lb="Arabidopsis thaliana TDNA insertion lines"
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna_protocols.html"
BASE COUNT 11 a 3 c 4 g 15 t 1 others
ORIGIN
Query Match 76.9%; Score 10; DB 17; Length 34;
Best Local Similarity 76.9%; Pred. No. 4.3e+04;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AATCAANNANTTA 13
||||| |
Db 20 AATCAATATTTA 8
RESULT 5
A2439152 35 bp DNA linear GSS 03-OCT-2000
LOCUS 1M0229G08R Mouse 10kb plasmid UUCGM library Mus musculus genomic
DEFINITION clone UUCGM0229G08 R, plasmid sequence.
ACCESSION A2439152
VERSION A2439152.1 GI:10563165
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognath; Muridae; Murinae; Mus.
REFERENCE Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
1 (bases 1 to 35)
AUTHORS Islam,H., Longacre,S., Mahmood,M., Meenen,E., Pedersen,T., Relliy
M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausen,A.
and Wright,D., Weis,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
JOURNAL plasmid inserts
COMMENT Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0229 row: G column: 08
Seq primer: CACACAGGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 35.
FEATURES
source
Location/Qualifiers
1..35
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCGM0229G08"
/clone_1lb="Mouse 10kb plasmid UUCGM library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(<http://www.jax.org/resources/documents/dnares/>). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of pMD42 (g114732114[gb]/AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."
BASE COUNT 13 a 8 c 0 g 14 t
ORIGIN
Query Match 76.9%; Score 10; DB 17; Length 35;
Best Local Similarity 76.9%; Pred. No. 4.4e+04;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 AATCAANNANTTA 13
||||| |
Db 6 AATCAATCATTTA 18
RESULT 6
BH854728/c 36 bp DNA linear GSS 08-JUL-2002
LOCUS BH854728 36 bp DNA linear GSS 08-JUL-2002
DEFINITION SALK_088935.40.95.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_088935.40.95.x, DNA
sequence.
ACCESSION BH854728
VERSION BH854728
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadri nab
1 (bases 1 to 36)
AUTHORS 'C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.

TITLE Zimmerman, J. and Ecker, J.R.
A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome
JOURNAL Unpublished (2001)
COMMENT Contact: Joseph R. Ecker
The Salk Institute Genomic Analysis Laboratory (SIGNAL)
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: eckersalk.edu
This is single pass sequence recovered from the left border of TDNA. This sequence lies within 300 bases of the 5' end of AT5G3910.
Class: TDNA tagged.

FEATURES
source Location/Qualifiers
1..36
/organism="Arabidopsis thaliana"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_088935.40.95.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/cdna_protocols.html"

BASE COUNT
12 a 3 c 5 g 16 t

ORIGIN

Query Match
Best Local Similarity 76.9%; Score 10; DB 17; Length 36;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 AATCAANNANTTA 13
||||| | |||
Db 22 AATCAATATTTA 10

RESULT 7
BH854701/c
LOCUS SALK_088987.31.80.x Arabidopsis thaliana TDNA insertion lines
DEFINITION Arabidopsis thaliana genomic clone SALK_088987.31.80.x, DNA sequence.
ACCESSION BH854701
VERSION BH854701.1 GI:21704291
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 37)
Alonso, J.M., Leisner, T.J., Barajas, P., Chen, H., Cheuk, R., Gadriab, C., Jeske, A., Karnes, M., Kim, C.J., Parker, H., Prednis, L., Shinn, P., Zimmerman, J. and Ecker, J.R.
A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome
JOURNAL Unpublished (2001)
COMMENT Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGNAL)
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: eckersalk.edu
This is single pass sequence recovered from the left border of TDNA. This sequence lies within 300 bases of the 5' end of AT5G3910.
Class: TDNA tagged.

FEATURES
source Location/Qualifiers
1..37
/organism="Arabidopsis thaliana"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_087236.29.60.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/cdna_protocols.html"

BASE COUNT
12 a 3 c 7 g 15 t

ORIGIN

Query Match
76.9%; Score 10; DB 17; Length 37;

source 1..37
/organism="Arabidopsis thaliana"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_088987.31.80.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/cdna_protocols.html"

BASE COUNT
12 a 3 c 6 g 15 t

ORIGIN

Query Match
76.9%; Score 10; DB 17; Length 37;

RESULT 8
BH854911/c
LOCUS SALK_087236.29.60.x Arabidopsis thaliana TDNA insertion lines
DEFINITION Arabidopsis thaliana genomic clone SALK_087236.29.60.x, DNA sequence.
ACCESSION BH854911
VERSION BH854911.1 GI:21704501
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 37)
Alonso, J.M., Leisner, T.J., Barajas, P., Chen, H., Cheuk, R., Gadriab, C., Jeske, A., Karnes, M., Kim, C.J., Parker, H., Prednis, L., Shinn, P., Zimmerman, J. and Ecker, J.R.
A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome
JOURNAL Unpublished (2001)
COMMENT Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGNAL)
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: eckersalk.edu
This is single pass sequence recovered from the left border of TDNA. This sequence lies within 300 bases of the 5' end of AT5G3910.
Class: TDNA tagged.

FEATURES
source Location/Qualifiers
1..37
/organism="Arabidopsis thaliana"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_087236.29.60.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/cdna_protocols.html"

BASE COUNT
12 a 3 c 7 g 15 t

ORIGIN

Query Match
76.9%; Score 10; DB 17; Length 37;

Best Local Similarity 76.9%; Pred. No. 4.5e+04;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AATCAANNATTA 13
|||||
DB 23 AATCAATTAATTTA 11

RESULT 9

AZ359655/c 46 bp DNA linear GSS 02-OCT-2000
LOCUS 1M0102A08R Mouse 10kb plasmid UGCC1M library Mus musculus genomic
DEFINITION clone UGCC1M0102A08 R, DNA sequence.

ACCESSION AZ359655
VERSION AZ359655.1 GI:10473355
KEYWORDS GSS.

SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
AUTHORS 1 (bases 1 to 46)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10Kb
Plasmid Inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0102 row: A column: 08
Seq primer: CACACAGGAACACGCTATGACC
Class: plasmid ends
High quality sequence stop: 46.

FEATURES

source

1. 46
Location/Qualifiers

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UCC1M0102A08"
/clone_lib="Mouse 10kb plasmid UGCC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g11473211419b1AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
ORIGIN

9 a 5 c 7 g 25 t

Query Match

76.9%; Score 10; DB 17; Length 46;

Best Local Similarity 76.9%; Pred. No. 4.9e+04;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AATCAANNATTA 13
|||||
DB 32 AATCAATGAATTA 20

RESULT 10

AZ508716/c 46 bp DNA linear GSS 05-OCT-2000
LOCUS 1M0551114F Mouse 10kb plasmid UGCC1M library Mus musculus genomic
DEFINITION clone UGCC1M0551114 F, DNA sequence.

ACCESSION AZ508716
VERSION AZ508716.1 GI:10690032
KEYWORDS GSS.

SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
AUTHORS 1 (bases 1 to 46)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10Kb
Plasmid Inserts
JOURNAL Unpublished (2000)
COMMENT Contact: Robert B. Weiss
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0351 row: I column: 14
Seq primer: CGTGTGAACACGCGCCAGT
Class: plasmid ends
High quality sequence stop: 46.

FEATURES

source

1. 46
Location/Qualifiers

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UCC1M0551114"
/clone_lib="Mouse 10kb plasmid UGCC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel
electrophoresis. Vector DNA was prepared from a derivative
of PMD42 (g11473211419b1AF129072.1), a copy-number
inducible derivative of plasmid R1. The vector was ligated
with adaptors complementary to the insert adaptors and
purified. The sheared, adapted mouse DNA was annealed to
adapted vector DNA, and transformed into
chemically-competent E. coli XL10-Gold (Stratagene) cells
and selected for ampicillin resistance."

BASE COUNT
ORIGIN

17 a 7 c 5 g 17 t

Query Match

76.9%; Score 10; DB 17; Length 46;

Best Local Similarity 76.9%; Pred. No. 4.9e+04;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 AATCAANNANTTA 13
||||| |
Db 41 AATCAAGACTTTA 29

RESULT 11

AZ481968

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

BASE COUNT

ORIGIN

Query Match

76.9%; Score 10; DB 17; Length 47;

12 a 3 c 7 g 25 t

10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 AATCAANNANTTA 13

||||| |

41 AATCAAGACTTTA 29

47 bp DNA linear GSS 04-OCT-2000

1M0306110R Mouse 10kb plasmid UUGC1M library Mus musculus genomic

clone UUGC1M0306110 R, DNA sequence.

AZ481968

GI:10643033

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 47)

Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C.,

Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly,

M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A.

and Wright, D., Weis, R.

Mouse whole genome scaffolding with paired end reads from 10kb

plasmid inserts

Unpublished (2000)

Contact: Robert B. Weiss

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT

84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert length: 10000 Std Error: 0.00

Plate: 0306 row: 1 column: 10

Seq primer: CACACAGGAAACAGCTATGACC

Class: plasmid ends

High quality sequence stop: 47.

Location/Qualifiers

1. 47

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC1M0306110"

/clone_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PMD42nv; Purified genomic DNA from M.

musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of PMD42 (g14732114[dbAF129072.1]), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to

adapted vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

BASE COUNT

ORIGIN

Query Match

76.9%; Score 10; DB 17; Length 47;

12 a 3 c 7 g 25 t

10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 AATCAANNANTTA 13

||||| |

41 AATCAAGACTTTA 29

49 bp mRNA linear EST 08-NOV-2001

AV856952

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

BASE COUNT

ORIGIN

Query Match

76.9%; Score 10; DB 10; Length 49;

10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 AATCAANNANTTA 13

||||| |

23 AATCAAGACTTTA 11

49 bp DNA linear GSS 02-MAY-2002

SAUK_049225 Arabidopsis thaliana rDNA insertion lines Arabidopsis

thaliana genomic clone SAUK_049225, DNA sequence.

BH810378

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

BASE COUNT

ORIGIN

Query Match

76.9%; Score 10; DB 10; Length 49;

10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 AATCAANNANTTA 13

||||| |

23 AATCAAGACTTTA 11

49 bp DNA linear GSS 02-MAY-2002

SAUK_049225 Arabidopsis thaliana rDNA insertion lines Arabidopsis

thaliana genomic clone SAUK_049225, DNA sequence.

BH810378

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

BASE COUNT

ORIGIN

Query Match

76.9%; Score 10; DB 10; Length 49;

10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 AATCAANNANTTA 13

||||| |

23 AATCAAGACTTTA 11

49 bp DNA linear GSS 02-MAY-2002

SAUK_049225 Arabidopsis thaliana rDNA insertion lines Arabidopsis

thaliana genomic clone SAUK_049225, DNA sequence.

BH810378

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

BASE COUNT

ORIGIN

Query Match

76.9%; Score 10; DB 10; Length 49;

10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 AATCAANNANTTA 13

||||| |

23 AATCAAGACTTTA 11

49 bp DNA linear GSS 02-MAY-2002

SAUK_049225 Arabidopsis thaliana rDNA insertion lines Arabidopsis

thaliana genomic clone SAUK_049225, DNA sequence.

BH810378

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

BASE COUNT

ORIGIN

Query Match

76.9%; Score 10; DB 10; Length 49;

10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 AATCAANNANTTA 13

||||| |

23 AATCAAGACTTTA 11

49 bp DNA linear GSS 02-MAY-2002

SAUK_049225 Arabidopsis thaliana rDNA insertion lines Arabidopsis

thaliana genomic clone SAUK_049225, DNA sequence.

BH810378

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

BASE COUNT

ORIGIN

Query Match

76.9%; Score 10; DB 10; Length 49;

10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 AATCAANNANTTA 13

||||| |

23 AATCAAGACTTTA 11

49 bp DNA linear GSS 02-MAY-2002

SAUK_049225 Arabidopsis thaliana rDNA insertion lines Arabidopsis

thaliana genomic clone SAUK_049225, DNA sequence.

BH810378

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

BASE COUNT

ORIGIN

Query Match

76.9%; Score 10; DB 10; Length 49;

10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1 AATCAANNANTTA 13

||||| |

23 AATCAAGACTTTA 11

49 bp DNA linear GSS 02-MAY-2002

SAUK_049225 Arabidopsis thaliana rDNA insertion lines Arabidopsis

thaliana genomic clone SAUK_049225, DNA sequence.

BH810378

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

BASE COUNT

ORIGIN

Query Match

Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA. This sequence lies within an annotated intron of At3g04140.
Class: TDNA tagged.

FEATURES
source

1.49
/organism="Arabidopsis thaliana"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_049225"
/note="PCR was performed on Arabidopsis thaliana lines"
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/tdna_protocols.html"

BASE COUNT
21 a 5 c 3 g 20 t

ORIGIN

Query Match 76.9%; Score 10; DB 17; Length 49;
Best Local Similarity 76.9%; Pred. No. 5e+04;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AATCAANNATT 13
||||| |
Db 2 AATCAATACTTA 14

RESULT 14
BH861490/c 49 bp DNA linear GSS 05-AUG-2002
LOCUS SALK_081694 Arabidopsis thaliana TDNA insertion lines Arabidopsis
DEFINITION thaliana genomic clone SALK_081694, DNA sequence.
ACCESSION BH861490
VERSION BH861490.1 GI:22096816
KEYWORDS CSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 49)
Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
, C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P.,
Zimmerman,J. and Ecker,J.R.
A sequence-indexed library of insertion mutations in the
Arabidopsis Genome
Unpublished (2001)
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA. This sequence lies within 300 bases of the 5' end of
At5g3910.
Class: TDNA tagged.

FEATURES
source

1.49
/organism="Arabidopsis thaliana"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_081694"
/note="PCR was performed on Arabidopsis thaliana lines"

BASE COUNT
15 a 4 c 6 g 24 t

Query Match 76.9%; Score 10; DB 17; Length 49;
Best Local Similarity 76.9%; Pred. No. 5e+04;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AATCAANNATT 13
||||| |
Db 35 AATCAATATTTA 23

RESULT 15
A0257483

LOCUS A0257483 51 bp mRNA linear EST 25-APR-2002
DEFINITION A0257483 3'-directed mouse cDNA library Mus musculus cDNA clone
BED0010659 3', mRNA sequence.
ACCESSION A0257483
VERSION A0257483.1 GI:20322147
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 51)
Kato,K. and Matoba,R.
Generation of expressed sequence tags from mouse brain
Unpublished (2002)
Contact: Kikuya Kato
Graduate School of Biological Sciences
Nara Institute of Science and Technology
8916-5 Takayama, Ikoma, Nara 630-0101, Japan
Tel: 81-743-72-5581
Fax: 81-743-72-5589
Email: kkatob@is.tnara.ac.jp.
URL: <http://love2.alst.nara.ac.jp/BED/index.html>.

FEATURES
source

1.51
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="BED0010659"
/clone_lib="3'-directed mouse cDNA library"
/tissue-type="brain"
/note="Vector: pGEM-T-easy"
BASE COUNT 18 a 9 c 17 t

Query Match 76.9%; Score 10; DB 9; Length 51;
Best Local Similarity 76.9%; Pred. No. 5.1e+04;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AATCAANNATT 13
||||| |
Db 20 AATCAGACTTA 32

Search completed: December 26, 2002, 14:02:52
Job time : 1979 secs

THIS PAGE BLANK (USPTO)